NUREG/CR-1278

Handbook of Human Reliability Analysis With Emphasis on Nuclear Power Plant Applications

Draft Report for Interim Use and Comment

Prepared by
A. D. Swain, H. E. Guttmann

Sandia Laboratories
Albuquerque, NM 97185

October 1980

Prepared for
Probabilistic Analysis Staff
Office of Nuclear Regulatory Research
U.S. Nuclear Regulatory Commission
Washington, D.C. 20555
NRC FIN No. A1188

8010300543

HANDBOOK OF HUMAN RELIABILITY ANALYSIS WITH EMPHASIS ON NUCLEAR POWER PLANT APPLICATIONS

ABSTRACT

The purpose of this handbook is to aid qualified persons in evaluating the effects of human error on the availability of engineered safety features and systems in nuclear power plants. The handbook expands the human error analysis presented in WASH-1400 and includes principles of human behavior and ergonomics, analytical procedures, mathematical models, and human error probabilities derived from related performance measures and experience. The derived probabilities should be adequate to determine the relative merits of different configurations of equipment, procedures, and operating practices within a plant, and for gross comparisons among plants. Limitations of the handbook and cautions to be observed in its use are explicitly stated.

TABLE OF CONTENTS

TABLE OF CONTENTS

TABLE OF CONTENTS

TABLE OF CONTENTS

TABLE OF CONTENTS

TABLE OF CONTENTS

TABLE OF CONTENTS

TABLE OF CONTENTS

TABLE OF CONTENTS

LIST OF TABLES

LIST OF TABLES

LIST OF TABLES

LIST OF FIGURES

LIST OF FIGURES

FOREWORD

The idea for this handbook came from Dr. W. E. Vesely, Division of Systems and Reliability Research,* Office of Nuclear Regulatory Research, U. S. Nuclear Regulatory Commission. Dr. Vesely, the current project monitor, and Dr. M. C. Cullingford, the previous project monitor, also of the Division of Systems and Reliability Research, provided guidance and encouragement for this work. Dr. Vesely's technical contributions aided us materially. Much of Chapter 6, "Unavailability," was written by Dr. Vesely.

Special thanks are due Mr. Jens Rasmussen, Electronics Branch, Risø National Laboratory, Denmark, for his critical reviews, and to Mme. Annick Carnino, Atomic Energy Commission, France, and Mssrs. A. E. Green and A. J. Bourne, National Centre of Systems Reliability, United Kingdom Atomic Energy Authority, England, for their comments and encouragement. Mr. J. M. Wiesen, Manager, Reliability Analysis Department, and Drs. Richard R. Prairie and Robert G. Easterling, Statistics, Computing, and Human Factors Division, all of Sandia National Laboratories, made substantial contributions to the quantitative aspects of the handbook. A special paper by Dr. Easterling is included as an appendix to Chapter 7, "Dependence." We express our thanks to Barbara J. Bell, of Sandia's human factors group for her technical review of the entire draft.

Thanks are also due to several participants of the 1979 IEEE Standards Workshop on Human Factors and Nuclear Safety (Schmall, 1980) who reviewed early drafts of some chapters and an advance copy of the handbook. The present draft has benfitted materially from their technical reviews.

---

*Formerly Probabilistic Analysis Staff.

As noted on the title page, this is a draft for public review. Please send your comments to A. D. Swain, Division 1223, Sandia National Laboratories, Albuquerque, NM 87185. Your comments will be considered for the final version of this handbook to be prepared in 1981. Please forward your comments by March 1, 1981.

One comment we anticipate is the need for a workbook to present a step-by-step procedure for conducting a human reliability analysis of operations in nuclear power plants. We are preparing such a workbook for use by teams of reliability analysts without human factors training who are conducting risk assessments of nuclear power plants (NPPs) under the NRC's Interim Reliability Evaluation Program (IREP). The IREP will include an evaluation of a selected sample of operating NPPs in the U.S.

The most significant differences between this copy and the advance copy of March 1980 are as follows:

1. The handbook has undergone a detailed review to improve the comprehensibility of the technical material and to correct errors. However, we can almost guarantee that some errors remain.

2. There were several changes in uncertainty bounds, but only relatively minor changes in the estimates of nominal human error probabilities (HEPs).

3. The number of references has been increased, including references to basic tasks with experimental data on which some of our statements about human performance are based.

4. Chapter 6, "Unavailability," has been revised, including new examples, and its presentation has been simplified.

5. An appendix has been added to Chapter 11, "Unannunciated Displays," to illustrate calculation of mean and median numbers of trials to detection of a deviant display.

6. Chapter 13, "Valving Operations," has been completely reorganized, but very few of the HEPs have been changed, and these only slightly.

7. A section on arithmetic calculations has been added to Chapter 14, "Task Procedures."

8. The section on valves in Chapter 20, "Derived Human Error Probabilities and Related Performance Shaping Factors," has been changed to reflect changes in Chapter 13, and values in the last section (Graphic Representation of HEPs) have been changed to reflect the changes in nominal HEPs and their uncertainty bounds.

9. In Chapter 21, "Examples and Case Studies," the calculations in the section on Bounding Analysis have been changed to reflect the changes in HEPs and uncertainty bounds made throughout the handbook.

PART I. BASIC CONCEPTS

PART I. BASIC CONCEPTS

CHAPTER 1. INTRODUCTION

Purpose of Handbook

The primary purpose of this handbook is to present methods, models, and estimated human error probabilities (HEPs)* to enable competent analysts to make quantitative or qualitative assessments of occurrences of human errors in nuclear power plants (NPPs) that affect the availability or operational reliability of engineered safety systems and components. A second purpose of the handbook is to show the user how to recognize error-likely equipment design, operating policies and practices, written procedures, and other human factors problems so that improvements can be considered. Many studies have indicated that in complex man-machine systems human error has often been the overriding contribution to actual or potential system failures (e.g., Shapero et al, 1960; Meister, 1962; and Meister and Rabideau, 1965). Analyses of NPP operations indicate that NPPs are not exceptions to this general finding (WASH-1400; Rasmussen, 1976; and Rasmussen and Taylor, 1976). Finally, accidents such as those at Brown's Ferry and Three Mile Island (TMI) clearly show that humans have acted not only as accident initiators and accident propagators, but also as accident mitigators in NPPs.

It is our intent that this handbook assist utilities to evaluate the role of operating personnel in existing power plants, enable designers of

---

*Certain terms are defined in the glossary. These terms are underlined the first time they appear in the text. The meanings of all abbreviations are listed in the abbreviations section at the end of the handbook.

future plants to avoid major human factors problems, and provide a quantitative base for the assessment of human errors in NPP safety, effectiveness, and efficiency.

Although the handbook is oriented towards engineered safety features (ESFs), the models, procedures, and estimated HEPs are relevant to all aspects of NPP design and operation where there is an interaction of people with plant systems and equipment. Most of the material in this handbook is also applicable to human factors aspects of other large process plants; e.g., chemical plants, oil refineries, and other power-generating plants.

Relationship of Handbook to WASH-1400

Sandia National Laboratories personnel were involved in the reliability analyses performed in WASH-1400. (The authors of this handbook were the human reliability analysts for that study.) The human reliability assessments appear in various volumes of WASH-1400. Section 6.1, "Human Reliability Analysis," in Appendix III, Failure Data, describes in general terms how the estimates of HEPs for various system safety tasks were derived and incorporated into the system fault trees.

Since WASH-1400 presents only summaries of the human error analyses, it is sometimes difficult for readers to understand how the various HEPs were developed. To utilize human reliability principles more fully in plant design and operations, more information is needed than that given in WASH-1400. Particularly, information is needed that can be applied to specific problems in NPPs. In this handbook we define the concepts involved, the data employed, and the calculations used in applying human error analyses to system evaluations (of reliability or availability) in NPPs. It is intended that the methodology and estimated HEPs presented should apply to NPPs in general, not only to *light water reactors* (LWRs), the subject matter of WASH-1400.

Limitations of the Handbook

The state-of-the-art in human reliability analysis is barely beyond its infancy. Until recently, many system reliability or system safety analysts did not attempt to quantify the effects of human performance. Even today, numerous system reliability and system safety analyses omit human error analyses, or they make unrealistic simplifying assumptions concerning the probabilistic nature of human error. Neither of these approaches is satisfactory as either can lead to erroneous and possibly dangerous conclusions in risk assessment studies. Experience in military, space, and commercial man-machine systems indicates that the human has a major role in both accident propagation and mitigation. Despite limitations in the coverage and accuracy of human performance estimates, use of the models and estimated HEPs in this handbook can lead to realistic risk assessments and reliability analysis in general.

Human performance is difficult to predict because of its variability. Any given operator in an NPP differs from all other operators, and will frequently show remarkable variability in behavior from day to day and from moment to moment. The human performs more different functions, in more different ways, under more different conditions than any other single element in a system. He has more interfaces, he accepts a greater variety of inputs, he provides a greater variety of outputs, and the possible relationships between his inputs and outputs are even more varied.

Despite this variability, it is possible to predict, within error bounds, the reliability of a human involved in a task for which he is adequately trained. It is even possible to estimate grossly the variability among differently trained and experienced humans performing that task. The reader must bear in mind that the error bounds may be broad. Therefore, the user of this handbook should not expect his estimates of error probabilities to be precise.

The reader must also understand that, if inexperienced in analyzing human performance, his estimates could have broader error bounds than those stated. The most frequent mistake made by those who are not experienced in human performance analysis is to ignore the various types of interrelationships that exist among operators (including supervisors), between operators and equipment, and between operators and operating conditions, including the various formal and informal feedback paths that modify human behavior. Another mistake is to assume that people will always do what they are told to do (either by oral directions, by written instructions, or by plant policy). If either of these mistakes is made, the analyst's estimates of HEPs are likely to be too optimistic, and he will ignore certain behaviors that could have serious impact on the system. If the user is aware of the difficulty of estimating failure probabilities of equipment but believes that human behavior is easier to understand and predict, he, too, will be subject to unjustified optimism.

The more the user of this handbook knows about human behavior in systems, especially in nuclear power systems, the more accurate his identification of human events and estimation of HEPs are likely to be. There is no substitute for experience in any systematic endeavor, especially in one with as large an element of subjectivity as human reliability analysis. In our opinion, the best human reliability analyses will be those performed by teams of experts in various areas, specifically including highly qualified human factors personnel.

Jens Rasmussen, a Danish authority in the human reliability field, describes a handbook like this as analogous to a handbook for surgery to be used by a ship's captain: quite a bit can be done with such a handbook, but some things require an expert -- a surgeon in one case, a human reliability analyst in the other. This handbook will often be misused by the naive analyst so that human errors will be assessed as insignificant when they actually are significant. This view is pessimistic, but there are no easy ways to estimate the effects of human errors on NPP safety, and the best estimates will be only approximate.

On the more optimistic side, approximations are adequate for most human reliability estimates. If the user realizes that estimates of HEPs are made with a sizeable range of uncertainty and that his final estimate of human influence is based on such uncertainty, he will be less likely to err in his evaluations.

Another limitation of the handbook is that we were unable to develop models and estimate HEPs for <u>all</u> NPP tasks. Our emphasis is on the kinds of tasks that we addressed in the WASH-1400 study -- calibration, maintenance, and selected control room tasks related to the availability of ESFs. We have not studied certain other tasks such as the use of computer systems and video readouts, or those involved in refueling, plant security, plant evacuation, emergency communications, and plant chemistry. Finally, the HEPs and models are based on studies and observations in existing, conventional LWR plants, such as Surrey, Peachbottom, Dresden, Zion, Calvert Cliffs, and San Onofre, which provide commercial power. Some newer plants may incorporate human factors improvements which could make some of our estimates too pessimistic. Despite these limitations, the user will be able to apply much of the material to tasks not specifically considered in this handbook, since there can be considerable similarity in human factors aspects of different plants despite differerences in equipment and other engineering aspects.

The scarcity of objective and quantitative data on human performance in NPPs is a serious limitation. Most of the HEPs in this handbook are what we call derived data. In some cases, they are extrapolations from related (sometimes only marginally related) performance measures. In other cases the HEPs represent our best judgment based on our experience in complex systems (including NPPs) and on our background in experimental and engineering psychology. This necessity of relying on judgment is a regrettable state of affairs, but a start needs to be made, and this handbook is a first step towards what is really needed -- a large data bank of human performance information directly related to NPP tasks.

A final point, which some may consider a limitation, is that the handbook does not deal with malevolent behavior. This is a handbook about human errors made by people who intend to do the correct thing but sometimes fail in this intent. Malevolent behavior is not due to error: it is deliberate behavior calculated to produce a harmful effect.

Organization of the Handbook

In addition to five major parts divided into chapters, this handbook consists of prefatory sections, references, equations, a glossary, and abbreviations. The detailed table of contents serves as an index. Part I, "Basic Concepts," consists of this chapter (1) and two others. Chapter 2 presents the basic definitions of terms, including a categorization of types of errors that one can expect in NPPs, fossil fuel power plants, or any man-machine system. Chapter 3 presents some philosophy and guidelines for viewing the human in a system context, including a discussion of factors that influence human performance in a system. Chapter 3 also lists principles of good human factors design, along with NPP examples of conformance with and deviation from these principles.

Part II, "Method for Analysis and Quantification of Human Performance," consists of three chapters. Chapter 4 presents the analytical methods used to identify the tasks and task elements to be used in the human reliability model. Chapter 5 presents the human reliability model, the related probability tree diagramming, and a general procedure for performing a human reliability analysis. The relationship of human reliability analysis to system reliability studies and to other types of event and fault trees is discussed. Chapter 6 briefly describes the use of the HEPs to estimate the unavailability of systems and components due to human error.

Part III, "Human Performance Models," consists of Chapters 7 through 18, which present models developed from available experimental literature, interviews with and observations of NPP personnel in the U.S. and in Europe, and the experience of the authors. The human performance models address those time relationships in operator behavior that are important in estimating recovery factors either for human-initiated failures or for the detection of nonnormal plant situations. The models are presented as mathematical statements, with uncertainty bounds when appropriate. These models involve considerable extrapolation from available data and experience, and should be regarded as hypotheses. It is hoped that these models will be subjected to rigorous testing in laboratory and plant settings so that appropriate modifications can be made.

Part IV, "An Interim Human Performance Data Bank," consists of two chapters. Chapter 19 discusses the sources of the HEPs and the models in the handbook. Chapter 20 consolidates the HEPs from the preceding chapters for convenient reference.

Part V, "Application of the Handbook and Concluding Comments," consists of two chapters. Chapter 21 presents some case studies to illustrate task analysis and the application of the human performance models, HEPs, and the human reliability technique to NPPs. Chapter 22 presents concluding comments, including an assessment of current human reliability analysis techniques and apparent trends.

It is intended that there will be a companion volume to this handbook entitled Human Performance Data Related to Nuclear Power Plant Operations that will include a compendium of human performance data from the files of the Sandia National Laboratories' Human Factors Group. These data will be catalogued according to a taxonomy of human actions related to NPP tasks. Also to be included are three data banks of derived human performance data. These are the AIR Data Store (Munger et al, 1962; Payne and Altman, 1962; Payne et al, 1962; and Smith and Payne, 1962), the Bunker-Ramo Data Bank (Meister, 1967), and the Aerojet-General Data Bank (Irwin et al, 1964a and b; and Meister, 1964). These data banks, though developed in the 1960s, are still useful. They are out of print;

hence the decision to include them in the companion volume. The companion volume will also describe the kinds of human performance data that would be most useful for human reliability analyses of NPP operations. A number of suggestions for collecting and collating such data will be included.

How to Use this Handbook

In writing this handbook, it was not possible to follow a step-by-step sequence from the first to the last page. The subject matter is complex and the procedure highly iterative. We suggest that the user read through the entire volume, paying particular attention to the mechanics of human reliability analysis described in Part II and illustrated in examples throughout the handbook, especially in Chapter 21. The user should work out some of these examples for himself because unless he develops skill in probability tree diagramming, especially in its representation of the conditional probabilities of events, the likelihood of his performing a satisfactory human reliability analysis will not be very high.

The many examples of how human behavior and performance are estimated under various situations in NPPs constitute a "scenario-oriented" approach. We hope that the user can develop a "feel" for how humans behave, since precise approaches do not as yet exist for modeling human behavior in all its complexities and with all its interactions. The handbook presents basic principles, guidelines, a reasonable amount of modeling, a set of human performance data, and numerous examples to assist the user in performing a human reliability analysis. In this sense, this document is not like the usual handbook in which one can look up some set of data and apply it directly to a problem. Because of the diverse backgrounds of those interested in human reliability in NPPs, much of the handbook is tutorial. Some of the human factors information will seem elementary to those with a background in human factors technology, and some of the information on reliability technology will seem elementary to those with a background in that area. We have tried to integrate the information so that practitioners of both technologies will have sufficient guidance to function as part of a team of human reliability analy ts.

Once the user has developed a facility for probability tree diagramming and understands the limitations and rationale for the estimated HEPs in the handbook, he should find that the summary tables and information in Chapter 20 will be all that he need consult for solving most human reliability analysis problems for NPP operations.

CHAPTER 2. EXPLANATION OF SOME BASIC TERMS

Although the glossary defines all specialized terms used in the handbook, this chapter elaborates on some that require additional discussion. The order of presentation of the terms in this chapter was chosen to facilitate their development:

page

Human Engineering - Human Factors Engineering - Human Factors - Ergonomics

All of the above terms describe a discipline concerned with designing machines, operations, and work environments so that they match human capacities and limitations (Chapanis, 1965, p 8). The first three terms are used most widely in the United States. The last term is used primarily in other countries, but is being used more frequently in the United States.

People working in the human factors area are often called human factors specialists or engineering psychologists. In Europe, the term ergonomists is used. In this handbook, these terms are interchangeable.

Man-Machine System and Interfaces

The term man-machine system denotes a system in which people have a monitoring and/or control function. The term man is used in the generic sense. The term man-machine interface refers to points of interaction between people and components in a system. Thus, a display, a control, or any other item a person observes or operates is a man-machine interface.

Human Reliability

Evans (1976) notes that the popular definitions of reliability and availability are as follows:

> Reliability is the probability of successful performance of a mission.
>
> Availability is the probability that the system or component is available for use when needed.

Meister (1966) defines human reliability as "the probability that a job or task will successfully be completed by personnel at any required stage in system operation within a required minimum time (if the time requirement exists)." We borrow from Evans and Meister to define human reliability as the probability of successful performance of the human activities necessary for either a reliable or an available system.* We include in this definition the probability that a system-required human act, task, or job will be completed successfully within a required time period, as well as the probability that no extraneous human actions detrimental to system reliability or availability will be performed. This definition is in keeping with Green and Bourne (1972, p 1) who note that "... the quality of a man's performance and also the time at which or in which he performs may be a measure of his reliability in association with any particular task."

Human Reliability Analysis

Human reliability analysis is a method by which human reliability is estimated. The method commonly used in solving practical human reliability problems is the one described in this handbook in Chapter 5. In carrying out a human reliability analysis it is necessary to identify those human actions that can have an impact on system reliability or availability. The most common application of human reliability analysis is the evaluation of human acts required in a system context. The consideration of extraneous human actions is also important. The human in a system may not only fail to do what he is supposed to do, or fail to do it correctly, but he may also do something extra that could degrade the system. The latter is the weak link in human reliability analysis. It is not possible to anticipate all undesirable extraneous human actions. The best anyone can do is identify those actions with the greatest potential

---

*In other applicatons, other measures of human performance (e.g., interval or ordinal numbers) can be used to define human reliability, but in this handbook we use probabilities only.

for degrading system reliability and availability. The assignment of probability estimates to extraneous actions is very difficult and uncertain. Often the best we can do is estimate very broad ranges of HEPs that we believe include the true probability in question.

Human Reliability Model

A model of a system is an abstraction that reproduces symbolically (simulates) the way in which the system functions operationally (Chapanis, 1961). In this handbook the term human reliability model denotes a schematic representation or abstraction of human events and related system events and their interactions in a man-machine system. When probability values are assigned to the elements in the model, the resulting mathematical expressions provide estimates of the probabilities of achieving (or not achieving) certain combinations of events in the system.

Human Error

The 1975 issue of Webster's New Collegiate Dictionary defines an error as "an act involving an unintentional deviation from truth or accuracy." This definition is close to the one employed in this handbook. Several of the other definitions in the dictionary connote blame or fault on the part of the person who makes the error. In this handbook we define human error as any member of a set of human actions that exceeds some limit of acceptability (Rigby, 1970). Thus, an error is merely an out-of-tolerance action, where the limits of tolerable performance are defined by the system. There is no connotation of blame or fault. If human errors are analyzed with the same objectivity as are other out-of-tolerance system components, successful corrective or preventive action can be taken. If errors are judged to be due to characteristics of the person who makes them, the resultant emotional atmosphere will mitigate against a rational determination of appropriate corrective action.

It is convenient to distinguish between errors whose primary causal factors are related to the design of the work situation and errors whose primary causal factors are related to some human characteristic. The term situation-caused error (SCE) is used for the former; human-caused error (HCE) is used for the latter. Ergonomists have long recognized that most errors in a well-defined work situation such as an NPP are due to SCEs and that relatively few are due to HCEs. This is hardly surprising. People who make an inordinately large number of HCEs don't last very long on a job.

The thrust of this handbook, therefore, is toward presenting methods and techniques to identify and quantify the effects of SCEs. An approach that emphasized HCEs would not be cost-effective. Although people do sometimes deliberately fail to do the right thing, the most effective approach to error analysis is to look for the SCEs involved. If maintenance personnel failed to follow a written procedure step by step, the HCE approach would result in blaming the individuals for the failure, citing them for such vague characteristics as "poor motivation" or "carelessness." In an approach that emphasized SCEs one would want to know why the written procedures were not used. In actual applications, it is nearly always found that the procedures are inconvenient to use because they are not written in accordance with ergonomics principles. On the other hand, if it were found that one individual persistently refused to follow procedures that were well-written while other maintenance personnel did use the procedures in the intended manner, errors made by that individual would be classified as HCEs. The offending person could be told, with justification, that correct use of the procedures was a necessary condition for continued employment.

In the human factors field, the SCE approach to error analysis is commonly used. The primary causal factors behind most human errors in a well-structured work situation such as an NPP are more closely related to such system elements as operating procedures, equipment design, and management practices than to the individual characteristics of trained personnel. Recognizing this fact, we have named this technique the Work Situation Approach (WSA) (Swain, 1969c, 1980a).

By convention, the definition of human error normally excludes malevolent behavior but does include intentional errors. These latter occur when the operator intends to perform some act that is incorrect but that he believes to be correct or to represent a superior method of performance. An erroneous belief in the correctness of a procedure often results from some misinterpretation or failure to understand an order. The operator's belief that his way is better than the prescribed way can result in a deliberate violation of standard operating rules. Examples of such errors include (1) not using written procedures in the intended manner because they are poorly designed (an SCE) or because the operator is truly lazy (an HCE), (2) deliberately loosening some equipment tolerances (setpoints) because management has harshly penalized operating personnel for shutting the reactor down when it turned out to have been unnecessary (an SCE for the operator, an HCE for management), and (3) venting low-radioactive containment pressure to the atmosphere because the operator is not willing to wait for the automatic safety features to respond to the increasing containment pressure (an SCE if training is inadequate, an HCE if the operator is merely impatient). Intentional errors also include out-of-tolerance behavior arising from a disorganized emotional state; e.g., an operator who blindly operates switches on a panel when a major accident occurs. Such behavior is rare, but can occur.

Most errors are unintentional -- the error just happens; it was not intended. Examples include (1) spilling coffee over the control board (which might be classified as a combination HCE, violation of rules, and SCE, if there is no convenient, safe place for the operator to place his coffee cup, (2) inadvertent tripping of the reactor when an operator sits on the edge of the control panel (an SCE because this behavior should have been anticipated and guarded against, but also an HCE because this behavior is clearly inappropriate and operators know this), and (3) activating an incorrect control because the intended control is located nearby and the labels for the two controls are very similar in appearance (an SCE).

It is important not to equate an error with its consequences. Sometimes a single human error will result in undesirable system consequences. Often it will not; for example, the spilled cup of coffee lands on the floor instead of on a control panel. However, this error had the potential to cause damage to the control panel. It was just luck that nothing untoward occurred--except for a messy floor. We would call this error a no-cost error. No-cost errors are important--they are signals that preventive action should be taken promptly to reduce the possibility of serious consequences.

Categories of Human Error

A person in a system can make an error if he does something incorrectly, fails to do something he should, or fails to do something in time. It is convenient to think of five major categories of human error:

(1) An error of omission - a person fails to perform the task or part of the task (e.g., a step).

(2) An error of commission - a person performs the task or step incorrectly.

(3) An extraneous act - a person introduces some task or step that should not have been performed.

(4) A sequential error - a person performs some task or step out of sequence.

(5) A time error - a person fails to perform the task or step within the allotted time, either too early or too late.

The latter three categories are errors of commission but are listed separately because their causal factors are frequently different.

From a systems point of view, any one of the above behaviors is considered an error only when it reduces or has the potential for reducing system reliability, system safety, or the likelihood that some other system success criterion will be met. Obviously, a person in a system performs many extraneous acts; e.g., smoking a cigarette, scratching his nose, and the like. In a system context, these behaviors are not considered errors unless they have potential for degrading the system in some manner. Sometimes an error can result in an undesirable consequence (i.e., an unrecovered error), but generally, just by chance or because of recovery factors in a well-designed system, no serious loss to the system will occur (i.e., the error is a recovered error).

Human Error Probability

In the human reliability technique described in this handbook, the basic measure of human performance is the human error probability (HEP). The HEP is the probability that when a given task is performed an error will occur. There are many ways to estimate the HEP; some are statistical and some are nonstatistical (i.e., judgmental). We will use the term human error probability to represent any estimate.

The most useful information for human reliability analysis is actuarial data; i.e., HEPs which consist of the known number of errors of a given type divided by the number of opportunities for that error to occur. This is expressed as:

$$\text{HEP} = \frac{\text{number of errors of a given type}}{\text{number of opportunities for the error}}$$

Our HEP is the measure defined by Green and Bourne (1972, p 22) as $P_f = N/n$, where P_f is the proportionate number of failures, N is the total number of failures, and n is the total number of events. In our earlier reports, including Section 6.1 of WASH-1400, we used the term human error rate (HER) interchangeably with human error probability. Although our use of HER was correct, it is not used in this handbook to avoid confusion by those who believe that the term rate must be associated with time.

If a data-based estimate is not available, an estimate derived from information on similar tasks can be used. Similarity is judged in terms of the correspondence of behavioral variables. Two physically dissimilar items of equipment might be similar in terms of the human behaviors involved in their operation, calibration, or maintenance. Therefore, an observed HEP for one of these items of equipment might be used as the estimate of the HEP for the same task on other items of equipment.

The probabilities most often used in human reliability analysis can be classified as demand probabilities; that is, the probabilities that given human actions will be performed and performed correctly when required. If time limitations are imposed on the performance of a task, one probability of interest is the probability that the task will be completed correctly within the allotted time. If required, the HEP per hour can be obtained. For most availability calculations, the interest is in the probability of at least one error (for a given task) per hour. In availability estimates, the HEP per hour is estimated even though the task may be performed with a frequency of much less than once per hour. Some sample calculations are presented in Chapter 6.

The reliability of a task; i.e., the probability of its successful performance, is generally expressed as: 1 - HEP. Thus, when we speak of the reliability of performance of a human task, we are thinking of the probability of successful performance per demand. When we speak of the error probability, we mean the probability of unsuccessful performance per demand, or task unreliability, which is 1 minus task reliability. The terms, human error probability, human failure probability, or task failure probability are often used interchangeably with human unreliability. (The same can be said for human success probability, task success probability, and human reliability.)

Basic, Conditional, and Joint Probabilities

Three types of probability are important in performing an analysis. These are the basic human error probability (BHEP), the conditional human error probability (CHEP), and the joint human error probability (JHEP).

BHEP is the probability of a human error on a task which is considered as an isolated entity, unaffected by any other task. If the task is the first in a series of tasks, there is no ambiguity in this definition. If the task in question is not the first task and its outcome may be dependent upon the outcome of other tasks, the BHEP would be that probability conjectured to exist if no other tasks were involved.

CHEP is the probability of human error on a specific task given failure, or success, on some other task. Two tasks are independent if the CHEP is the same regardless of whether success or failure occurred on the other task; otherwise, they are dependent.

JHEP is the probability of human error on all tasks which must be performed correctly to achieve some end result. This is the probability of most interest in reliability work and is determined by using both BHEPs and CHEPs.

Uncertainty Bounds

When an estimated HEP for a task or a human action is presented in the handbook, it is usually followed in parentheses by a range expressed as a lower and upper HEP bound. These error bounds (called uncertainty bounds) reflect uncertainty that arises from two sources. One source is associated with variability due to people and conditions. The other source is the uncertainty in our assessment of the error probabilities. Thus, the expression .01 (.002 to .02) means that our best estimate of the HEP (i.e., nominal HEP) is .01 and that we believe it is unlikely that the HEP would in any case be lower than .002 or higher than .02. By unlikely we mean that there is only about a 10% chance that an HEP could be lower than .002 or higher than .02. That is, on a distribution of HEPs that represents our assessment of the relative likelihood of various values of the HEP, .002 represents the lower 5th percentile and .02 represents the upper 5th percentile. It is obvious that other analysts could propose other values for point estimates and bounds.

Unavailability because of Human Error

Availability was defined as the probability that a system is available for use when needed. Its converse, unavailability, is one minus availability. In NPPs, any errors of operation, maintenance, or calibration can result in the unavailability of some safety-related system or component for some period of time. This unavailability continues until someone discovers that the system or component is not operative, or until its condition causes other changes to the plant that lead to the discovery. In addition, other system events can cause some ESF to be unavailable, and this unavailability may be displayed on some meter or result in some other visible change in the plant. Plant personnel then have the opportunity to note this change and take steps to restore the unavailable ESF to its normal operating condition.

The role of human performance in the unavailability of ESFs is discussed in Chapter 6, "Unavailability."

Variability of Human Performance

As mentioned in Chapter 1, variability is a characteristic of human performance. Humans never do anything exactly the same way twice. Of course, most human variability is of no consequence to a system. For example, the fact that Operator A takes 2.3 seconds to respond to an annunciator one time and 3.1 seconds the next time usually does not matter to the system. As long as the variability of the human performance is within limits defined as acceptable for system operations, no error has occurred. It is only when the response is outside the system-specified human tolerance limits that an error has occurred. The narrower the limits, the more likely it is that an error will occur. Thus, human variability can contribute to human error, and the larger the variability the larger the HEP will be for most situations.

We define three classes of human error: random, systematic, and sporadic. In the paragraphs which follow we treat these errors separately; in the real world a given human output may include one, two, or (infrequently) all three categories.

Random errors are out-of-tolerance actions that follow no predictable pattern but occur when the variability of behavior results in performance that is beyond system-acceptable variability. An example of random error can be illustrated by the case of a length of pipe that is measured many times by an individual, in which the values resulting from many measurements are distributed about an average according to some probability distribution. The average of all measurements may or may not be the true length of the pipe, but the measurement average is not far off the true average. If the natural variability of the individual is such that some of his measurements are outside the variability of measurements accepted by the system as tolerable, these measurements would be considered random errors.

Random error probabilities can be reduced either by reducing the variability of performance or by enlarging the acceptable tolerances. Random errors can occur when the skill of the operator is not adequate so that, even with generous system tolerances, his variability is such that some of his performances are unacceptable. Random errors can also occur even with the small variability in performance of skilled operators if the system tolerance limits are extremely tight or if the operators cannot control some significant factor such as high stress. Normally, more random errors occur among novices than among skilled operators in a well-defined operational setting such as an NPP. Under high levels of stress, greater variability in performance can be expected and the random error probability will increase, even among skilled personnel. (Stress is discussed in Chapters 3 and 17.)

Systematic errors are out-of-tolerance actions characterized by a dispersion pattern offset from a desired norm; that is, there is a consistant bias. This would occur with the pipe measurements if the measuring device were miscalibrated. If the bias is large, the error will be large even if the variability is small.

Systematic errors can occur as a result of a bias in conjunction with random variation in that one or more of the person's acts may be outside of the established tolerance limits. For example, a calibration technician may deliberately intend to adjust certain setpoints on the high side to minimize the possibility of a trip. We expect a certain small amount of variability in the adjustment he makes. Since he is biasing his adjustments in one direction, occasionally his natural variability may result in a setting that is actually beyond the safe limit for a setpoint.

As in the above example, systematic errors are most likely when an operator is concerned with only one limit of a range of acceptability. This often occurs in inspection tasks, and is called inspector flinching. For example, the inspector may "fudge" his performance to be on the safe side. Thus, a safety-oriented pressure vessel inspector may reject acceptable welds that are close to the minimum limits even though they are within the range of acceptability. In effect, his bias has raised the mean value of acceptable welds. Conversely, a production-oriented inspector may accept welds that are slightly below the minimium. His bias would displace the mean value of acceptable welds in the opposite direction.

Systematic errors can also occur when an operator is not given adequate feedback on his performance, as when he is trying to adjust some parameter but his controls and displays have an imperfect relationship to this parameter. Such a relationship occurs when there is a lag between a control adjustment and the parameter being controlled; for example, as when rod adjustment is done manually and there are feedback lags.

Biases can exist in tools and instructions and can also result from the operator's personality, training, or experience. Providing an operator with specific and timely feedback of his performance is usually the best way of controlling systematic errors that are not built into the system.

Sporadic errors are infrequent actions that are outside the tolerance limits, despite small variability in performance, as when a skilled marksman occasionally fires a "wild" shot. This outlier is a sporadic error; its occurrence is a surprise to all concerned. Those errors made by skilled operators are generally sporadic errors. Therefore, when they occur, it is usually pointless to insist that the skilled operator be more careful (the HCE approach). Such errors often occur when a person is distracted; e.g., a step in a maintenance procedure is skipped because the maintenance technician is interrupted by a phone call and then resumes the task at the wrong place. Since sporadic errors occur infrequently and are often made by well-trained and experienced workers, it is difficult to determine the causes. The best prevention is to collect data from large numbers of operators, to analyze the conditions under which sporadic errors were made, and to correct the conditons.

Human Tolerance Limits

Since human error is defined as an out-of-tolerance response, limits must be placed on human responses to keep the variability of human behavior within acceptable tolerances. These limits, referred to as human tolerance limits, are employed in NPPs, as in any work situation. Several types of tolerance limits are used. The following list arranges tolerance limits from the most effective to the least effective:

Barrier Limits physically prevent or limit unacceptable performance.

For example, stops on a hoist prevent the object being raised from crashing into the roof, even if the operator keeps his finger on the FAST UP button. A plastic guard over a TRIP button reduces the chances of unintentional (spurious) trips by guarding against inadvertent activation of the switch.

Fixed Limits are clearly and permanently established limits. An example is the use of green (acceptable) and red (unacceptable) patches on instruments. Detents on switches make it difficult to leave the switch between functional positions. Lines on the plant floor define safe passageways.

Empirical (or Measurement) Limits are checked by observation or measurement during or after performance. For example, a meter indication is checked to see that it is within tolerances. A setpoint is adjusted until a meter reads a desired value. A sample of containment atmosphere is used to measure radioactivity level. A hole is drilled and then measured to see if it is within tolerance.

Reference Limits are standards to be compared with an output in time of doubt. Examples are: samples of just barely good welds and just barely bad welds that are provided for verification purposes.

Caution Limits are given by warnings, signs, or other indications. These are not among the most effective human tolerance limits because they are often not present while an action is being performed, or, if present, they are a familiar part of the worker's environment, and no longer are attention-getting signals.

Conventional Limits are those instilled by training or custom, but they may not be otherwise reinforced in the work situation. For example, reliance may be placed on "good shop practices" such as: "Put tools away," "Cut away from your body," or "Don't lean against the control board."

Forensic Limits are argumentative (subject to debate) and are often defined after some incident has occurred to assess blame. For example, when an accident occurs, it is often argued that the person involved in the accident should have known that his performance was likely to have caused the accident, and it is concluded, after the fact, that he exceeded the limits of safe performance. This reasoning is usually circular, and cannot be considered an effective human tolerance limit.

The above tolerance limits are listed in decreasing order of effectiveness, but this listing also represents very closely the order of increasing frequency of application. Although Conventional Limits are probably the most commonly used, the frequency of human error can generally be reduced most effectively by using tolerance limits from the top of the above list.

Performance Shaping Factors

Many factors affect human performance in a complex man-machine system such as a nuclear power plant. Some of these performance shaping factors are external to the person in the system, and some are internal. Human reliability is affected by the entire work environment, especially the equipment design and the written or oral work procedures. The individual himself brings to the job certain skills, motivations, and expectations that influence his performance. Psychological and physiological stresses result from a work environment in which the demands placed on the operator by the system do not comform to his capabilities and limitations.

To perform a human reliability analysis, an analyst must understand those performance shaping factors (PSFs) that are most relevant and influential in the jobs studied. Chapter 3 discusses several of the PSFs that influence the reliability of nuclear power plant personnel.

Dependence, Independence, and Coupling

Dependence between two tasks refers to the situation in which the probability of failure on one task is influenced by whether a success or failure occurred on the other task. In WASH-1400 the term coupling was used, whereas in this handbook we use the term dependence. Complete dependence between two tasks means that if failure occurs on one, failure will occur on the other with certainty. Similarly, if success occurs on one task, success will occur on the other. (It is possible that the 100% correlation may be negative, but for the usual situations, the correlation will be positive as stated.)

Zero dependence between two tasks means that the probability of failure or success on one task is the same regardless of whether failure or success occurred on the other. In human reliability analysis the assumption of zero dependnece, although it may be unjustified, is often assumed for convenience in situations in which the analysis is not materially affected.

To illustrate the concept of dependence, consider a situation in which two tasks (task "A" and task "B") are performed by the same person in succession. Let task "A" be the calibration of one gauge and task "B" be the calibration of a second dissimilar gauge. Assume, ideally, that human error data are available on a very large number of results on the calibration of the two gauges. Further, assume that the BHEP on both tasks is .05.

Complete dependence would pertain if the probability of failure on task "B" given failure on task "A" was 1 and the probability of failure on task "B" given success on task "A" was zero. Zero dependence would pertain if the probability of failure on task "B" was .05 regardless of whether failure or success had occurred on task "A." Some level of intermediate dependence would pertain if, for example, the probability of failure on task "B" given failure on task "A" was .20 while the probability of failure on task "B" given success on task "A" was .04.

In this handbook, five levels of dependence are used: zero dependence (ZD) (i.e., complete independence), low dependence (LD), moderate dependence (MD), high dependence (HD), and complete dependence (CD). These are discussed in detail in Chapter 7.

Human Error Consequences and Recovery Factors

As noted earlier, errors may not result in any serious consequences to a system. In WASH-1400, it was shown that most human errors would not materially reduce the availability of ESFs. Some errors may be no-cost errors; for example, an error that causes no damage or does not lower the availability of any ESF. For example, an incorrect switch is manipulated, but the operator immediately realizes his error and returns the switch to its proper positon before it has an adverse impact on the system.

Other errors could adversely affect some component or part of an ESF, but the adverse impact is prevented or compensated for by other components or systems, or even by other human actions taken at a later time. We call these preventive or compensatory factors recovery factors. Human redundancy is a recovery factor: a calibration technician makes an error, but a companion checking his work catches the error and it is corrected. Such an error is also a no-cost error.

Error-Likely Situations and People

Some work situations are obviously error prone (error likely) in the sense that the ergonomics are so poor that errors are likely to occur; hence, error-likely situations (ELSs). ELSs involve demands on humans that are not compatible with their capabilities and limitations. In the United States if a toggle switch is installed so that the up position is OFF, errors are likely because in this country we expect the opposite arrangement. If one is required to read and remember a 7-digit number for even a few seconds, while he also has some other task to do, errors in recall are likely.

Sometimes a person is characterized as error-likely. In an NPP, a truly error-likely person would soon be recognized and would be retrained, reassigned, or discharged. Although chronic error likeliness in people qualified, trained, and experienced in a job is not common, we are all error-likely from time to time. Anyone emotionally upset is usually more likely to make errors. If one is fatigued from unusually long hours of work or has not had enough sleep, certain types of errors are relatively likely. Error likeliness in people who have had adequate training in a job is usually temporary.

Accident-Prone Situations and People

Accident-proneness is a special case of error-likeliness. The accident-prone situation (APS) is one that fosters human errors likely to result in injury to people or damage to equipment and facilities. An accident-prone person is one who has statistically "more than his share" of accidents when compared with others having the same degree of exposure to opportunities for the same types of accidents.

A familiar example of an APS is a slippery floor. For example, a leak in a condensor system in a <u>boiling water reactor</u> (BWR) wet the floor, increasing the chances of someone's slipping while walking across it. When this temporary problem was cleared up, the APS no longer existed.

Accident-proneness in individuals is not a very useful concept, as it has at times been used to justify an HCE approach to solving safety problems. Although there are people who seem to be chronically accident-prone (for example, the class of male drivers in the United States under 25 years of age), studies of work situations show that accident-prone people in industry are rare. Carefully controlled studies show that accident-proneness in people, when it does occur, is most often due to temporary conditions such as illness or emotional disturbances.

In the early 1900s, the concept of the accident-prone individual in industry arose in part because faulty statistical analyses were used that did not incorporate concepts of statistical significance. Subsequent analyses of these early data showed that certain individuals were stigmatized as accident-prone when the number of accidents they experienced was not significantly greater than the number expected due to chance alone (Mintz and Blum, 1961). Even when chance can be ruled out, it may be found that people seeming to have "more than their expected share" of accidents are those who have the greatest exposure to the risk of accidents.

Taking all of the above factors into consideration, most modern industrial safety specialists conclude that it is more cost-effective to look for APSs than to look for accident-prone people. The emphasis in this handbook is on techniques for identifying APSs and ELSs and for estimating their potential impact on the availability of ESFs.

Populational Stereotypes and Expectances

A populational stereotype is the way in which members of a population expect things to behave especially with respect to directional movements. In the U.S. we expect the "UP" position of a light switch to turn the lights on. In Europe, the opposite populational stereotype holds.

Populational expectance is a type of populational stereotype that refers specifically to displays which do not involve directional movement. For example, in the U.S. the letters "H" and "C" on water faucets stand for "hot" and "cold." In Mexico, a different populational expectance holds for the letter "C" -- it stands for "caliente," which means hot.

Any design that violates strong populational stereotypes or expectances will result in a relatively high HEP. Even with extensive training, it is difficult to change a populational stereotype completely. Under stress, we tend to revert to our populational stereotypes.

Stressors and Stress (Physiological and Psychological)

Montaigne, a French essayist in the late 1500s, noted, "Men under stress are fools, and fool themselves." Although he was probably thinking primarily about psychological stressors, we can think of stressors as any external or internal forces that cause bodily or mental tension, or stress. Stress is the human response to a stressor. Thus, in an NPP there can be physiological stressors such as fatigue, discomfort, constriction of movement, or high temperature, as well as psychological stressors such as task speed, distractions, monotonous work, threats from supervisors, and emergency situations.

Chapters 3 and 17 present discussions of stress, including its facilitative and disruptive effects.

Types of Nuclear Power Plant Tasks

The handbook includes data and procedures relevant to many tasks in existing NPPs. Tasks of special interest, because of their potential impact on the availability of ESFs, are control room operations, preventive and corrective maintenance procedures, and calibration procedures.

Normal control room operations (including startup, shutdown, power level control, and the operator's participation in calibration and testing) are also of interest, especially to the extent that these operations affect ESF availability. Control room operations in response to transients and after other abnormal conditions (p 2-26) are of special interest since human error frequencies can be inflated under the stressful conditions involved.

Also of special interest are calibration and maintenance procedures that involve the adjustment of setpoints and the removal and restoration of ESF components after maintenance. Uncaught human errors here may result in the unavailability of ESFs for relatively long periods of time.

Task Taxonomy

There have been several attempts to develop a useful task taxonomy (Chambers, 1969, Fleishman et al, 1968, 1970). While these classifications of human tasks are useful for psychological research, they are not very useful for human reliability analysis because they refer primarily to human variables. The approach taken here is to categorize the data on human behavior in terms of combinations of equipment and task variables, a classification scheme developed earlier (Swain, 1956). This categorization is readily seen in Chapter 20 in which all of the human reliability data in the handbook are summarized. For example, when one is performing a human reliability analysis of NPP operations, the usual procedure is to provide estimated HEPs for tasks identified in a system fault tree analysis. Thus, the fault tree analysts identify a particular man-machine interface which has important implications for, say, the availability of some ESF. This interface might be a valve which may be placed in the wrong position for maintenance or may not be restored to its normal position after maintenance. Our primary search term is "valves." Secondary search terms are related to type of valves (manual or motor-operated) and human behavior terms such as "change" and "restore." Additional search terms are stated in terms of PSFs such as "level of stress," "skill level," "type of feedback to operator," and so on.

At this time we have not worked out a detailed search scheme (taxonomy) that would enable the user to choose key words and identify quickly the tasks and HEPs most relevant to a specific analysis. This is subject matter for the planned companion volume mentioned on p 1-8 and titled Human Performance Data Related to Nuclear Power Plant Operations. For the time being, Chapter 20 and the detailed table of contents provide the best sets of search terms.

Task and Link Analysis

Chapter 4 presents an analytical procedure called task analysis. This technique is used to identify the relevant human elements in tasks and to identify the ELSs and APSs in these tasks.

Task analysis is an analytical process for determining the specific behaviors required of the human components in a man-machine system. It involves determining the detailed performance required of people and equipment, and the effects of environmental conditions, malfunctions, and other unexpected events on both. Within each task, behavioral steps are analyzed in terms of the perceptions, decisions, memory storage, and motor outputs required, as well as in terms of the expected errors.

The data from a task analysis can be used to establish equipment design criteria, operating sequences, written procedures, and requirements for selection and training of personnel. Task analysis is the most commonly used tool of the human factors specialist, and is required for a human reliability analysis.

Link analysis is a special form of task analysis. It depicts the pathways among different parts of a system that are generated by people walking about and communicating with each other.

Transients

A transient is any departure of some NPP function from established normal limits. Thus, it is an abnormal operating condition rather than a normal operating condition. It may be anticipated in the sense that it occurs with sufficient frequency that the operating staff is not unduly surprised when it occurs and a set of procedures is available for the response. It may be unanticipated in the sense that it is a rare or even unplanned-for event that generally evokes a reaction of surprise and disbelief.

Transients may occur with or without a reactor trip (or scram). Anticipated transients without scram (ATWS) can pose problems for operating personnel, who may have to interpret display readings and decide whether and when to trip the reactor manually.

LOCA

A loss-of-coolant accident (LOCA) is a transient in which there is a loss in the primary coolant system. The severity of a LOCA may range from the extreme (i.e., a guillotine break in one of the very large pipes) to a very small leak. The possibility of a LOCA has important implications for human reliability. In WASH-1400, the very conservative assumption was made that a "large LOCA" would involve the most severe conditions even though such a break has never occurred. It was judged that human reaction to such a condition would involve an extremely high level of psychological stress. For smaller LOCAs, initial human stress levels are assumed to be lower than for the very large LOCA. However, even small LOCAs can build to situations in which operating personnel manifest high stress.

Displays

A display is any instrument or device that presents information to any sense organ (visual, auditory, or other). In NPPs, displays are annunciated or unannunciated. Annunciated displays usually consist of panels of legend indicators (often called tiles) associated with an auditory signal. Unannunciated displays in NPPs include meters, digital readouts, chart recorders, graphs, indicator lights, computer printouts, and video presentations.

Manual Controls

Manual controls are those components with which the human enters his inputs to a system. Types of controls in NPPs include switches (rotary, toggle, and other), pushbuttons, levers, knobs, cranks, connectors, and tools. Manual controls may be continuous (e.g., a rod control lever) or discontinuous (i.e., discrete); e.g., a two-position rotary switch for a remotely operated motor-operated valve (MOV).

Continuous and Discontinuous Tasks

Manual control tasks may be either continuous or discontinuous (i.e., discrete). Continuous tasks involve some sort of tracking activity in which the operator monitors some continuously changing situation. The control action in a continuous task can be either continuous (as in rod control) or discrete (as in stopping a pump when water level reaches some point). A discontinuous (or discrete) task is one in which each task element is a discrete step (e.g., calibration).

CHAPTER 3. PERFORMANCE SHAPING FACTORS

This chapter describes many of the factors that affect human reliability in NPP operations. The purpose is to outline some of the qualitative aspects of human performance as a background for the quantitative models presented in Part III. Because of the lack of data and the limitations in our experience, not all of the performance shaping factors in this chapter are represented in the quantitative models.

The Human as a System Component

A fundamental assumption in human reliability analysis is that the human can be treated analytically, much as are other components in a man-machine system. This assumption has led some to the conclusion that the human functions are like any other system component. Considering the greater variability of human performance, such a conclusion is invalid. Still, human failures can be studied objectively and quantified, as can any other component failure.

Figure 3-1 depicts the human as a component that receives inputs, acts on these inputs, and produces outputs. As shown, the human is part of a closed-loop system, since information about his outputs is fed back to his sense organs to become another input. Thus, the human is able to discriminate any significant difference between his actual output and the desired output. This is the classical representation of the human as a system component (often called a "black box").

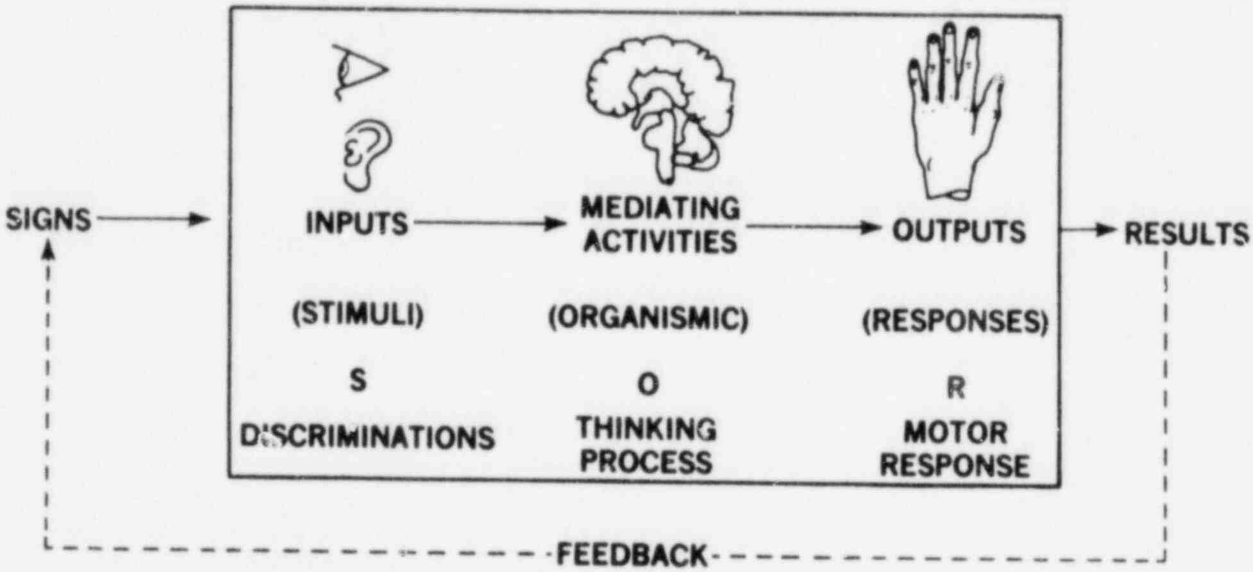


Figure 3-1. Human Operator in a Feedback System

The manner in which the human perceives, thinks about, and responds to the inputs he receives depends on what are called performance shaping factors (PSFs). The PSFs determine whether human performance will be highly reliable, highly unreliable, or at some level in between. In the present state of human reliability technology, it is not possible to assign weighting factors to each of these PSFs or to develop an equation for the relative influence of all of the PSFs on the performance of a given task. One reason for this limitation is that many of the PSFs interact, and the interaction effects are usually complex. Despite the current limitations in quantifying their effects, we must carefully consider all the relevant PSFs in evaluating task reliability.

In the human performance models presented in Part III, the estimated HEPs and uncertainty bounds are based in part on the judged influence of different combinations and levels of PSFs. If the user knows that the relevant PSFs are particularly good for a given situation, the calculated error probability should be assumed to lie closer to the low end of the range by some factor, say a factor of 2. Thus, if the calculated HEP is .001 (.0001 to .01), it could be modified to .0005 (.0001 to .01). The uncertainty bounds are unchanged. Conversely, if the PSFs are particularly unfavorable for human reliability, the HEP could be modified to .002 (.0001 to .01). Note that the factor of 2 is for example only; depending upon the quality of the PSFs, the user may elect to apply some other factor. If the PSFs are average for the industry, no change would be made.

Classes of Performance Shaping Factors

Table 3-1 presents the PSFs that must be evaluated when performing a human reliability analysis (Swain, 1967a, 1980b). In general, PSFs are divided into three classes: (1) the external PSFs, those outside the individual, (2) the internal PSFs, those that are a part of the individual himself, and (3) stresses and stressors.

External PSFs

In general, the external PSFs are those that define the work situations of the operators, technicians, maintenance personnel, engineers, clerks, and others who keep the NPP performing reliably and safely. The external PSFs fall into three general categories: Situational Characteristics, Task and Equipment Characteristics, and Job and Task Instructions. Situational Characteristics include PSFs that are often plant-wide in influence, or that cover many different jobs and tasks in the plant. Task and Equipment Characteristics include PSFs that are restricted to some given job or even to a task within a job. Job and Task Instructions include PSFs connected with the instructional materials used in jobs. Although instructions are really task characteristics, they are singled out because they represent an area in which a relatively small investment in time and money can result in substantial improvement in human reliability.

Internal PSFs

It is a truism that every human component in a man-machine system is a unique element. Each person comes to the job with certain skills, abilities, attitudes, and a host of other human attributes. It would be very convenient if one could select or develop standardized people for

Table 3-1. Performance Shaping Factors

EXTERNAL

Situational Characteristics

- Architectural Features
- Quality of Environment:
  - Temperature, Humidity, and Air Quality
  - Lighting
  - Noise and Vibration
  - Degree of General Cleanliness
- Work Hours/Work Breaks
- Availability/Adequacy of Special Equipment, Tools, and Supplies
- Manning Parameters
- Organizational Structure (e.g., Authority, Responsibility, Communication Channels)
- Actions by Supervisors, Co-workers, Union Representatives, and Regulatory Personnel
- Rewards, Recognition, Benefits

Job and Task Instructions

- Procedures Required (Written or not written)
- Written or Oral Communications
- Cautions and Warnings
- Work Methods
- Plant Policies (Shop Practices)

Task and Equipment Characteristics

- Perceptual Requirements
- Motor Requirements (Speed, Strength, Precision)
- Control-Display Relationships
- Anticipatory Requirements
- Interpretation
- Decision-Making
- Complexity (Information Load)
- Narrowness of Task
- Frequency and Repetitiveness
- Task Criticality
- Long- and Short-Term Memory
- Calculational Requirements
- Feedback (Knowledge of Results)
- Continuity (Discrete vs Continuous)
- Team Structure
- Man-Machine Interface Factors:
  - Design of Prime Equipment, Test Equipment, Manufacturing Equipment, Job Aids, Tools, Fixtures

STRESSORS

Psychological Stressors

- Suddenness of Onset
- Duration of Stress
- Task Speed
- Task Load
- High Jeopardy Risk
- Threats (of Failure, Loss of Job)
- Monotonous, Degrading, or Meaningless Work
- Long, Uneventful Vigilance Periods
- Conflicts of Motives about Job Performance
- Reinforcement Absent or Negative
- Sensory Deprivation
- Distractions (Noise, Glare, Movement, Flicker, Color)
- Inconsistent Cueing

Physiological Stressors

- Duration of Stress
- Fatigue
- Pain or Discomfort
- Hunger or Thirst
- Temperature Extremes
- Radiation
- G-Force Extremes
- Atmospheric Pressure Extremes
- Oxygen Insufficiency
- Vibration
- Movement Constriction
- Lack of Physical Exercise

INTERNAL

Organismic Factors

- Previous Training/Experience
- State of Current Practice or Skill
- Personality and Intelligence Variables
- Motivation and Attitudes
- Knowledge of Required Performance Standards
- Physical Condition
- Attitudes Based on Influence of Family and Other Outside Persons or Agencies
- Group Identifications

work in NPPs. Since this is not possible, an attempt is made to select workers who can develop acceptable levels of performance; the internal PSFs determine the potential level to which the individual can be developed. However, the methods employed to train the worker and to maintain or improve his proficiency are external PSFs established by the utility. As will be seen, there is ample room for improvement in the development and maintenance of skills, especially those skills related to coping with unusual events in NPPs.

Stressors

Because of its importance, stress, an internal PSF, is listed as a separate class of PSFs. Stress, either psychological or physiological, can arise when there is a mismatch between the external and the internal PSFs. For example, if the perceptual requirements of a task impose too many demands on a worker, performance will suffer because of excessive task loading, a psychological stressor. A well-designed man-machine system is one in which the task demands are consistent with the worker's capabilities, limitations, and needs. To the extent that this consistence is not achieved, human errors and adverse effects on motivation and attitudes can be expected.

Situational Characteristics

Architectural Features

Situational characteristics refer to PSFs that apply to more than one task. The first situational characteristic listed in Table 3-1 is architectural features. By this term we mean the general work area or areas. A control room, a room of equipment to be calibrated, the room housing the turbine and generator---all have certain architectural features that can affect human performance either favorably or adversely.

One familiar example in NPPs with both favorable and adverse impacts is the large control room which houses the operating panels for two or three reactors. One positive aspect is the fact that, in an emergency, centralized control is facilitated and it is more likely that sufficient qualified personnel will be available to cope with the emergency even in its earliest stages. A negative aspect is the generally high noise level, the excessive number of people present at times, and the possibility for confusion.

A problem in many control rooms arises from their sheer size. Operators often have to walk several yards or more to read some display. Because it is normal for people to avoid unnecessary effort, they try to read the displays from a disance and, consequently, make errors in their readings (Seminara et al, 1976, p 4-27).

Quality of the Working Environment

The next four situational characteristics in Table 3-1 (temperature, humidity, and air quality; noise and vibration; illumination; and degree of general cleanliness) refer to the quality of the environment surrounding the worker. NPPs generally provide a satisfactory environment, but, as noted earlier, there are exceptions regarding the noise level and an excessive number of people in control rooms. There are certain areas; e.g., the turbine room, where a high noise level is to be expected and ear protectors should be worn. However, a high noise level should not be tolerated in the control room as it can cause irritation and fatigue, which may result in errors. The problem appears to be more psychological than physiological, since noise levels in most control rooms are well below the levels specified under the Occupational Safety and Health Act (OSHA). In an informal, unpublished study, we took measurements with a sound level meter in various areas at Sandia National Laboratories where drafting and related support tasks are done. We found one consistent result: in any work area where the noise level was over 60 decibels, a significant number of occupants complained that the noise interfered with their concentration. Even though there are no physiological bases for these complaints, the occupants were irritated and distracted by the noise.

Methods for damping noise are well-known and will not be discussed here. In one Swedish NPP, carpeting of the control room floor was installed to reduce the noise level and thereby to reduce the physiological stressor of noise. There were also psychological implications inasmuch as this management action favorably impressed the operating personnel, who saw this as an example of the company's concern for their well-being.

Lighting for NPP tasks is often not adequate. For example, in some control rooms, glare is such a problem that the ambient illumination must be reduced, sometimes to the point that errors in reading displays may occur. In some areas in a plant, the lighting may be so poor that errors in reading valve labels are possible.

A special problem for certain NPP tasks is that of exposure to radioactivity and the requirement for wearing protective clothing when performing certain tasks. (We list radiation as a physiological stressor in Table 3-1, but it could also be listed as a psychological stressor or under quality of the working environment.) We have not studied this problem in detail, but interviews with operating personnel indicate that the clothing is uncomfortable and that a primary motivation of personnel in a "rad" environment is to get the job done as quickly as possible and "get out of there." This motivation mitigates against the highest levels of human reliability.

The degree of general cleanliness has a psychological impact quite apart from any impact on plant equipment. A generally dirty and untidy work environment may suggest to the worker that indifferent work performance is tolerable. Although we cannot quantify the influence of this PSF on human errors, it should be obvious that a clean and tidy work environment is more likely to affect human performance positively than the opposite kind of work environment. Although most NPPs maintain good housekeeping practices, there are substantial differences among plants.

Work Hours and Scheduling of Work Breaks

There have been many studies of human performance as a function of work hours and scheduling of work breaks. Most of these studies are relevant to production jobs or other work where the work rate is paced by the job. Much of the work in NPPs is self-paced except in response to unusual situations, in which case the workers are fully aroused and involved. The usual findings regarding work breaks are not applicable, since most of the NPP operators, technicians, and other personnel are relatively autonomous in performing their duties, and their efficiency is not likely to be improved by a rigid schedule of rest periods.

However, there is an important question about the impact of work hours that are longer than normal. This type of schedule occurs fairly often, as when someone stays on duty at the end of his shift to fill in for someone on the next shift, or during plant shutdown and startup operations that require the presence of certain key personnel throughout.

There is special interest in the effects of extra work hours on the probability that an operator will detect unannunciated indications of marginal or out-of-tolerance conditons, and on any decisions he may be required to make, especially those related to nonroutine conditons. In addition to the effects of the extra work hours, there can often be time stress on personnel during nonscheduled reactor shutdown activities. The time stress occurs because everyone is aware of the economic impact of unscheduled hours of plant unavailability. It is known from military studies that the effects of combined stress and fatigue are usually worse than the sum of their separate effects.

We made a literature search of the effects of work hours on visual detection and decision-making with implications for the scheduling of work hours for NPP personnel (Swain and Bell, 1980). This search clearly shows that the shape of the curve of performance vs hours for the evening and night shifts differs considerably from that for the day shift. This is especially true when personnel operate on a changing shift status. Most authorities ascribe this effect to the disruption of strong sleep habits.

In a study of Swedish NPP operations (Axelsson and Lundberg, 1975), operators on other than the day shift reported greater fatigue and difficulty in sleeping (including less hours of sleep). A review (Trumbull, 1966) notes the significant effects of losing even one night's sleep--which is not uncommon in NPPs. This and other studies (Bloom, 1961; Grandjean, 1968) indicate that as fatigue increases (especially fatigue due to loss of sleep), the detection of visual signals deteriorates markedly, choice behavior demands more time and exhibits more error, and reading rate decreases. Longer periods of sleep deprivation (from 50 to 60 hours) can lead to difficulties in mental arithmetic, inability to recall names and objects from recent conversation, and even to momentary hallucinations.

Other studies (Grant, 1971, on drivers of British Railways locomotives; McFarland, 1971, on aircraft pilots; and Grandjean et al, 1971, on Swiss air controllers) show that fatigue results in less attention to certain types of signals: personnel develop their own subjective standards of what is important, and as they become more fatigued, they ignore more signals. Bartlett (1951) called this the "phenomenon of the lowered subjective standard." Although these studies do not provide a direct, unequivocal answer to the question of how the performance of NPP personnel is related to hours on the job, some reasonable extrapolations can be made to the NPP situation.

As a practical matter, a person's subjective feeling of fatigue does not predict the quality of his performance. There are large individual differences in performance. Some people can perform well even when tired; others feel that they are doing well although they are not. Taking these findings into account, we offer the following suggestions for scheduling work hours beyond the normal work time:

(1) Under no circumstances should one work in an NPP for over 16 straight hours.

(2) There should be at least a 12-hour break between work periods for any individual.

(3) No one should work more than 60 hours a week.

(4) No one should work more than 100 hours in any two consecutive weeks.

(5) Ensure that working an extra shift would not mean that one has been without sleep for 24 hours.

Availability/Adequacy of Special Equipment/Tools and Supplies

The effect of the availability or adequacy of special equipment and tools and of supplies on job performance is often ignored in industry. Generally, it appears that these items are located and managed in a way to suit the toolcrib attendants rather than the people who will use them on the job. This reversal of priorities can adversely affect job performance. For example, if much effort is required to obtain certain tools, the worker will tend to make do with what he has, often to the detriment of his work.

A positive application of this PSF has been observed in some NPPs. Those who perform calibration tasks are provided with special carts with all the test equipment, tools, and procedures needed for the job. Because it is convenient to use the correct items, they are used.

Manning Parameters

The manning parameters refer to how many and what kinds of people are used to perform which types of jobs. Much has been written on this topic, but often one of the basic requirements for manning effectiveness is not implemented. Our experience in developing manning tables for technical jobs in military organizations indicates that unless manning is based on a thorough task analysis of all the tasks to be performed by an organization, there will be inadequacies in manning. People tend to define jobs in general terms rather than from a careful analysis of task demands. In NPPs there is a natural tendency to carry over the manning practices from fossil fuel power plants. While much of this carryover is probably relevant, we know of no systematic study of NPP manning that determines the extent to which NPPs have different manning needs.

We have observed differences between U.S. and European practices in the type of people assigned to control room functions. In the U.S., operating personnel generally are not college graduates. Often they are former fossil fuel plant operators or former operators of shipboard reactors from the U.S. Navy. In Great Britain most control room operators are graduate engineers or the equivalent. There are pluses and minuses to both practices. Using graduate engineers may have some advantages in coping with unusual events if their expertise is necessary. On the other hand, problems of boredom are accentuated with these personnel; daily task demands are not up to their capabilities and job expectations.

The use of technically trained noncollege personnel may involve less risk of job dissatisfaction, but there is a question of how well these personnel can cope with highly unusual events. We believe it is less a question of whether college training is necessary than of how job-relevant the training is. Certain obvious problems are addressed later in this chapter.

Another difference between U.S. and European manning is illustrated by the Swedish practice of having two separate job titles for operating personnel. For Swedish NPPs there are two types of control room operators called Reactor Engineers and Turbine Engineers. Neither is a graduate engineer but both have training and backgrounds comparable to that of U.S. operating personnel except for shipboard reactor experience. However, problems occur because the reactor engineers regard their jobs as more prestigious than those of the turbine engineers. Also, the usual path for promotion to shift supervisor is to advance from reactor engineer, not from turbine engineer. In some cases, a lack of cooperation has been observed between turbine and reactor engineers in which the former have caused reactor trips when they decide to perform certain activities without first checking with the reactor engineers. In the U.S., all control room personnel are Reactor Operators (ROs) of various levels: an unlicensed RO (often called an Auxiliary Operator), an RO, and a Senior RO. Because the pathway to advancement is clearcut in this case, the problems experienced among operating personnel in the Swedish plants have not been frequently reported in U.S. plants.

Organizational Structure and Actions by Others

A plant's organizational structure (e.g., authority, responsibility, and communication channels) and actions by supervisors, coworkers, union representatives, and regulatory personnel often fall into the sociological realm. Some of these actions have technical impact and are therefore appropriate to this handbook. Probably the most important area is administrative control of the status of safety systems and components.

When systems or components are removed from their normal operating status for maintenance or other purposes, proper restoration procedures may not be carried out, or certain necessary safety systems or components may not be available because of oversights or other errors. In NPPs considerable reliance is placed on human intervention to avoid such problems. Therefore, the kind of administrative control in a plant is very important. If administrative control is tight and the supervisors insist on rigid adherence to the control, it will be much less likely that some valve will be left in the wrong position or that some jumper will not be removed after maintenance. Chapters 13 and 15 detail the importance of sound administrative control in recovering from errors in system and component restoration.

Actions by regulatory personnel, especially by government regulatory personnel assigned to a given plant, will have a substantial effect on plant personnel and practices. For example, plant personnel will usually respond if an onsite federal inspector emphasizes good housekeeping. If the inspector frequently queries operators on how they would respond to hypothesized unusual events, the operators will tend to think about coping with the unusual. On the other hand, if the inspector spends most of his time checking paperwork, the plant will expend most of its effort to ensure acceptable paperwork. Thus, the personality of the inspector has a strong influence on the "personality" of the plant.

Rewards, Recognition, and Benefits

These PSFs have at least an indirect influence on the performance of technical jobs. Industrial and social psychology textbooks discuss the importance of these factors. Although they may have an impact on human reliability, they are outside the scope of this handbook.

Task and Equipment Characteristics

The task and equipment characteristics are PSFs that are task-specific. These are the factors evaluated in performing a task analysis, as described in Chapter 4. The following sections describe in general terms the task and equipment PSFs listed in Table 3-2, and subsequent sections deal with specific findings on these factors in typical LWRs.

Perceptual Requirements

The perceptual requirements of a task are determined by the task and equipment features that convey information to the personnel. Almost all of the perceptual requirements placed on the personnel are visual -- reading meters, charts, labels, etc. Auditory requirements are minor, requiring only the ability to hear and recognize various alarms. The crucial PSFs for displays are that they reliably convey the essential information to the user, and that the display attract his attention (if prompt response is required).

Motor Requirements

Motor requirements refer to control, adjustment, connecting, or other actions. Normally these are performed with the hands or feet. Speed of movement is rarely a problem in NPPs, although the strength required to operate a control can be a problem in some maintenance work in which "come-alongs" are needed to change the status of large manual valves. Generally, however, demands for muscular strength in NPP operations appear to be well within the capabilities of the fifth percentile female operator (Hertzberg, 1972). Precision of motor response is not a problem in NPPs, except for certain operations performed during refueling and rod manipulations. Most of the human factors problems in the design of controls in NPPs are related to the use of unnecessarily large control handles and to the poor location and labeling of controls.

Table 3-2.* Task and Equipment Characteristics

Perceptual requirements

Motor requirements (speed, strength, precision)

Control-display relationships

Anticipatory requirements

Interpretation

Decision-making

Complexity (information load)

Narrowness of task

Frequency and repetitiveness

Task criticality

Long- and short-term memory

Calculational requirements

Feedback (knowledge of results)

Continuity (discrete vs continuous)

Team structure

Man-machine interface factors:

Design of prime equipment, test equipment, manufacturing equipment, job aids, tools and handling equipment, fixtures

*From Table 3-1

Control-Display Relationships

The relationships between controls and displays refer to the compatibility of displayed information with the required movement of controls. Certain displays lead to certain expectances as to how a control should be moved. If the displays and their associated controls violate these expectances, the probability of human error will be very high, especially under stress (Fitts and Posner, 1967, p 25).

Anticipatory Requirements

These requirements refer to tasks in which a person has to be alert for some signal while performing another activity that also requires attention. Humans have definite limitations in this area. Man is essentially a single-channel mechanism (Fitts, 1951); that is, he can pay attention to only one thing at any instant in time. With practice he can rapidly switch his attention among several stimuli (Gabriel and Burrows, 1968), and it may appear that he is attending to several things simultaneously. Still, at any given moment he is attending to just one stimulus.

The skilled control room operator ordinarily has ample capacity for the tasks at hand. Under unusual conditions, however, he may be overloaded---that is, there may be so many visual and auditory signals competing for his attention that he is unable to divide his attention among them in the most effective manner. He will ignore some signals either because he doesn't perceive them or because he has had to select the more important ones. The performance model for annunciated displays reflects this effect of signal loading on operators (Chapter 10).

Interpretation Requirements

Interpretation requirements in NPPs are related to situations in which the presented information requires some mental processing. This means that the course of action implied by the information is not obvious, and interpretation of the data is required. The more interpretation that is required, the longer the response time and the greater the probability of error. Ease of interpretation (or minimization of the need for interpretation) is achieved by an optimum combination of operator training, design of displays, and design of written procedures. Less-than-adequate implementation of these PSFs will increase interpretation demands and thus increase human errors.

Decision-Making

The need to make decisions in a job can help keep the job interesting and challenging. Without any requirement for decision-making, most people become bored. Therefore, the best-designed jobs and tasks include the need for a person to use his decision-making ability. Errors occur when the information presented to the decision-maker does not adequately support the kinds of decisions he needs to make. If this happens, the person making an incorrect decision is likely to be faulted even though he has responded logically to the information he had. For example, in aircraft crash investigations a verdict of "pilot error" may be attributed to the pilot, without any attempt to uncover the factors that led to the wrong decision (Pressey, 1976).

Complexity (Information Load)

The complexity of a job is a function of the amount of information the worker must process and the amount of abstract reasoning or visualization required. Obviously, errors will be frequent if the job is too complex. Tasks in an NPP ordinarily are well within the capabilities of the workers. The experienced operator understands the working of the plant, and he processes information at a self-determined pace. However, in some plants the emergency procedures (such as those for LOCAs) introduce complexities that exceed the capabilities even of highly skilled operators. (For an example, see p 21-11.)

As the information load on an operator increases, a point may be reached at which the operator can no longer process information. As described later in the discussion on stress, he has several ways of compensating for overload, some of which can result in error. Some years ago, it was believed that the concept of information theory (as described by Wiener, 1948; and Shannon and Weaver, 1949; and extended to psychology by Attneave, 1959; Garner, 1962; Sheridan and Ferrel, 1974; and others) could be used to quantify information load. In application to practical situations, however, the utility of information theory has been disappointing. One reason is that the objective measure of information in tasks or task elements cannot assess the meaning that the individual himself attaches to each signal. For example, when shift change approaches, signals related to this event are more compelling to some operators than to others. Analyses of Licensee Event Reports (LERs) show that a disproportionate number of errors of oversight occurs within an hour in each direction of shift change.

Frequency and Repetitiveness

The frequency and repetitiveness of human actions are PSFs that have a dual relationship to human performance. Although the ability to perform reliably obviously increases with the frequency and repetitiveness of a task, highly repetitive tasks become boring and few workers will do their best on such jobs. The optimal tradeoff between reliability and boredom in the design of jobs remains unsolved in many industries, including nuclear power. Some calibration tasks, for example, involve so much repetition that it is a tribute to the patience of the technicians that so few errors are made.

At the opposite extreme, if a task that is performed infrequently must be performed correctly, frequent simulated practice is necessary (as with fire drills). NPP emergency procedures are required at unpredictable times and therefore should be practiced periodically.

Task Criticality

The criticality of a task as perceived by plant personnel will affect how much attention they devote to the task. This is especially true during periods of time stress or fatigue when one doesn't have the time or the energy to perform all tasks or task elements. A person at work will naturally devote more of his attention to those tasks or task elements that he considers most critical to the job. A person's perception of what is critical is obviously influenced by instruction from his supervisor and by the "old hands" with whom he works. For example, the operators' overriding concern with avoiding a solid pressurizer in the TMI accident is thought to have led to their failure to detect other, more important conditions (Kemeny, 1979).

Long- and Short-Term Memory Requirements

These requirements can often degrade human performance. Although long-term memory for facts is not one of man's outstanding capabilities, he does have a good capacity for remembering principles, strategies, contingencies, and other rules and their applications---provided that he has been properly taught and has a chance to practice them occasionally (Fitts, 1951).

Short-term memory is notoriously unreliable (Welford, 1976, p 101), yet many jobs place unrealistic demands for short-term memory on personnel. Short-term memory is less reliable than long-term memory because it lacks the long learning time and rehearsal associated with the latter. An example is looking up a new telephone number from the directory and keeping it in mind for the dialing process. Data from the Bell System indicate that under these circumstances a probability of error of 5% can be expected with the circular dial. If a person has to remember digits (or other meaningless information) for more than a few seconds, errors of 1 to 5 percent can be expected. For reliable performance, people must be supplied with properly designed forms on which to record data.

Calculational Requirements

The performance of even simple arithmetic accounts for many errors in technical work. This is in part because short-term memory is often involved, and the calculational aspect per se is associated with relatively high HEPs. In one study (Rigby and Edelman, 1968b), highly skilled inspectors routinely calculating familiar arithmetic problems involving simple addition, subtraction, and division made slightly more than 1% errors using paper and pencil. Our experience in industry shows that the relatively high frequency of calculational errors is always a surprise to those who design work operations. The typical "solution" for these problems is to tell those who make calculational errors to be more careful, an HCE approach that rarely works for long.

Feedback

Feedback, a term borrowed from engineering technology, refers to the knowledge of results that a person receives about the status or adequacy of his outputs. Without the feedback loop shown in Figure 3-1 a worker operates as an open-loop system and cannot perform complicated activities reliably.

It is recognized that feedback has motivational as well as informational aspects; the motivational aspects are outside the scope of this manual. The informational content of feedback provides a person with objective information on what he is supposed to do and whether he does it correctly, and with detailed information on when and how he failed to do it correctly. Feedback must always be timely. In some cases delays of even a few seconds can seriously degrade performance (Miller, 1953a, Section IV), especially for certain continuous tasks (described below).

Continuity (Discrete Versus Continuous)

As defined on page 2-28, this term refers to the extent to which a task is nondiscrete (i.e., involving continuous tracking activities), or the extent to which each task element is a discrete step (e.g., as in a typical calibration or maintenance procedure). In NPPs there are a number of continuous tasks, and supplying power is a continuous process. In this sense, NPPs are considered to be a special type of process industry, distinct from manufacturing industries in which items are manufactured or assembled. Edwards and Lees (1973 and 1974) describe the special human factors problems related to process industries. As they point out, process control is made difficult by the following PSFs:

(1) Several display/control variables interact.

(2) The process has a long time constant; i.e., a delay in feedback.

(3) The important variables often must be estimated rather than measured.

(4) Readings from widely separated instruments have to be collated, and the operator has to rely on short-term memory.

(5) The basic process is often difficult to visualize.

All of these PSFs are found in the typical NPP control room.

Team Structure

This term refers to the combinations of people performing work that must be done by two or more people. In this handbook we are not concerned with the sociological aspects of team makeup, only with the technical aspects of people working in teams. One major technical aspect has to do with the recovery factors made possible by having one person in a position to observe another's work either during or after completion of the work. The effects of these recovery factors are discussed more fully in Chapter 15.

Another technical aspect has to do with dividing tasks among different people to reduce task overload. An example is given (beginning p 21-14) in which the addition of a second operator to control rooms in certain NPPs is estimated to result in a substantial gain in human reliability for a task to be performed under abnormal conditions.

Man-Machine Interface Factors

This final task and equipment PSF is a catchall category covering all points of contact between the human and the hardware that were not covered in the other categories. It includes test equipment, tools, handling equipment, etc. Although these items all affect performance, they are not treated individually, since the general principles discussed in Chapters 11, 12, and 13 apply in most instances. Some of the problems in the design of the man-machine interface are described below, and provide some of the qualitative rationale for the human performance models and derived data presented in those chapters.

Some Ergonomics Problems in NPPs

This section describes some examples of poor ergonomics observed at various NPPs. The EPRI Review (Seminara et al, 1976), Raudenbush (1971, 1973) and Swain (1975) present detailed expositions of human factors problems in the design of control rooms. The examples in this section are based largely on our observations and interviews in several U.S. and European plants.

Some of the ergonomics problems observed were:

1. Poor design and layout of controls and displays
2. Poor labeling of controls and displays
3. Inadequate indications of plant status
4. Presentation of nonessential information
5. Inadequate labeling and status indications of valves

Poor Design and Layout of Controls and Displays

Typically, the design and layout of controls and displays in NPPs are not in accordance with recognized human factors practices. Controls and displays are not always grouped by function, nor are the functional groups always clearly delineated. At many plants the operators have applied colored tape to the panels in an attempt to delineate functional groups. (For a formal method of using tapes on existing panels, see Seminara, Eckert, et al, 1979; and Seminara et al, 1980.) At other plants the operators have been prevented from applying such markings by management policies that place aesthetics above the needs of the operators (Seminara et al, 1977). The lack of functional grouping may partially explain the tendency of operators to rely on one display instead of cross-checking with other instruments that might indicate whether the first display is operating properly. There are several LERs of operators relying on a chart recorder indication and performing inappropriate actions because of a stuck pen. In some incidents, operators apparently relied exclusively on one display when other displays would have indicated the appropriate actions required.

In many U.S. plants, the control room is roughly divided into three areas: Engineered Safety Feature panels, Reactor Control panels, and Balance of Plant panels. This does not mean that all the displays and controls used in coping with safety-related incidents are located on one panel. For example, in coping with a LOCA, the use of all three panel areas is required. The layout of controls and displays only partially supports the required operator activities. In one pressurized water reactor (PWR), rod control is hindered by the location of the necessary display far off to one side of its related control, forcing the operator to assume an awkward stance to read the display.

Several informal link analyses of control room activities have been performed (Swain, 1975; "Human Reliability Analysis," Section 6.1, WASH-1400, 1975; Seminara et al, 1977; and Seminara, Eckert, et al, 1979). It is clear that the layout of NPP control rooms does not match the operating needs of control room personnel very well. Figure 3-2 shows a formal link analysis for a BWR control room operator for a 23-minute period, and Figure 3-3 shows an expanded link analysis for an 8-hour shift for the same operator. The report from which these figures are taken (Axelsson and Lundberg, 1975) notes that the layout of the control room makes for considerable travel time and some reluctance on the part of operators to perform all checking functions on a timely basis.

Perhaps the most serious deviation from accepted ergonomics practices in the design of NPPs is the use of mirror-imaging of panels for a two-reactor control room. This practice consists of reversing the layout of the displays and controls from one set of panels to the other. (Fortunately, the mirror-imaging does not go all the way and reverse the components within a display or control.) Mirror-imaging aggravates the problems of inadequate panel layouts. Even highly experienced operators report moments of confusion in this kind of work situation, though they are fully trained and experienced on both layouts.

Poor Labeling of Controls and Displays

The labels and legends used in NPPs are not always clear. In some plants, labels on controls and displays are taken from construction

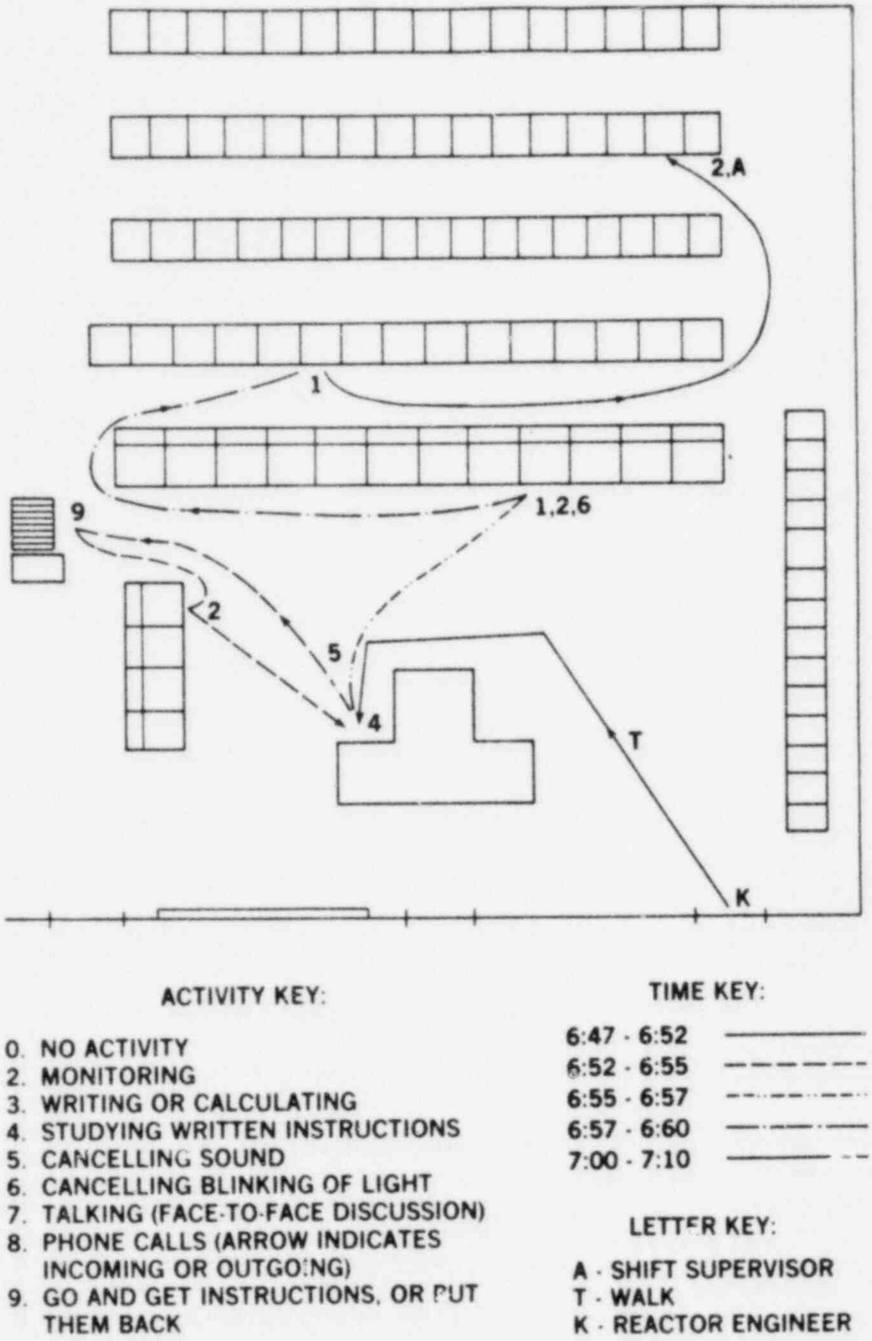
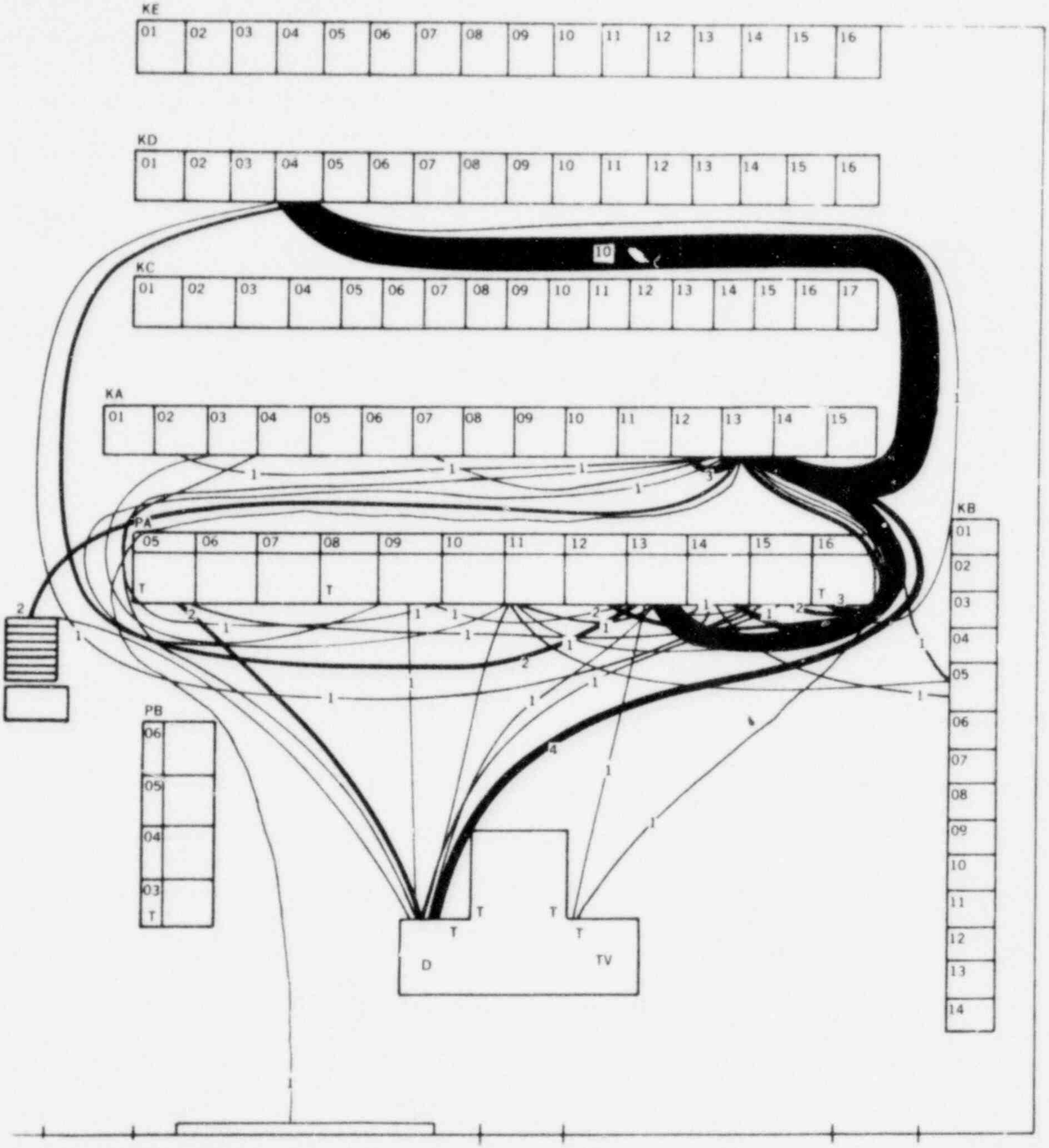


Figure 3-2. A 23-Minute Link Analysis of Reactor Engineer Activities in a Swedish BWR

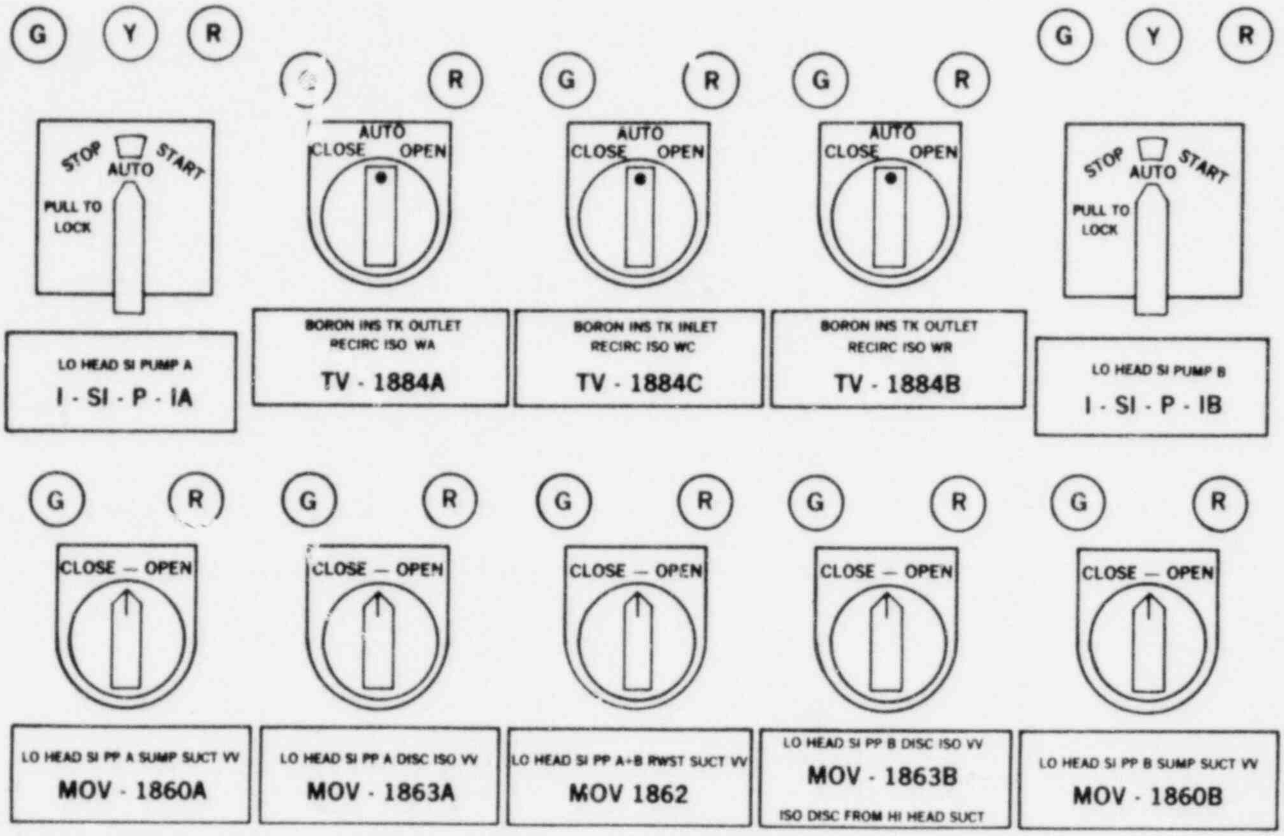


Key. The thickness of lines increases with the number of trips. The numbers 1 through 10 represent a rough numerical equivalent to the thickness of the lines.

Figure 3-3. An 8-hour Link Analysis of Reactor Engineer Activities in a Swedish BWR drawings with no attempt to develop labels useful to operating personnel. For example, consider the labels of the switches in the ESF portion of a panel shown in Figure 3-4. The labels of these switches describe their functions, which is helpful to the operator. However, the numbers differ so little that confusion among them is quite possible. The five switches immediately above these two rows are numbered: MOV-1864A, MOV-1885A, MOV-1885C, MOV-1885B, and MOV-1864B. Obviously, the numbers assigned to the switches provide no cues as to their locations on the panels.

In other plants, the possibilities for confusion and difficulties in locating controls and displays relevant to safety activities are even greater. For example, Figure 3-5 represents the lower ESF panel at one PWR. Note the labels and how difficult it would be to find the correct switch if it were not used frequently. We observed a highly skilled operator go through the LOCA procedures on the dynamic simulator which had panels identical to these NPP panels. Even this skilled operator had trouble locating certain switches required to cope with the simulated LOCA, and he described at length the greater problems experienced by less-skilled operators.

These difficulties prompted us to devise location aids similar to the matrix used on road maps (Swain, 1975). Figure 3-6 shows how this location aid would be applied to the panel shown in Figure 3-5. Figure 3-7 shows the same scheme for the upper ESF panel, and Figure 3-8 shows the location aid applied to an annunciator (ANN) panel. (This annunciator panel is labeled 1; the others would be labeled 2, 3, etc.) The five colors chosen can be discriminated even by persons with poor color vision (Baker and Grether, 1954). They are yellow (33538), blue (35231), white (37875), black (37038), and gray (36173). The 5-digit numbers



Key: The letters G, Y, and R refer to indicator lamps with, respectively, green, yellow, and red filters.

Figure 3-4. MOV Switches on part of the ESF Panel at the PWR used in the WASH-1400 Study
(The sketch is based on Figure III 6-2 from WASH-1400.)

Problem: Quickly Find Switch 8923B

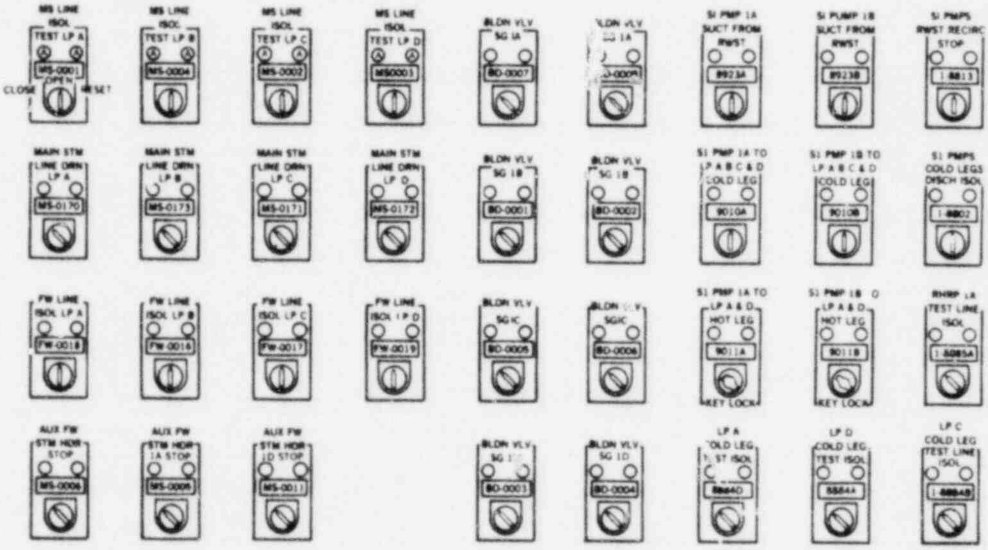
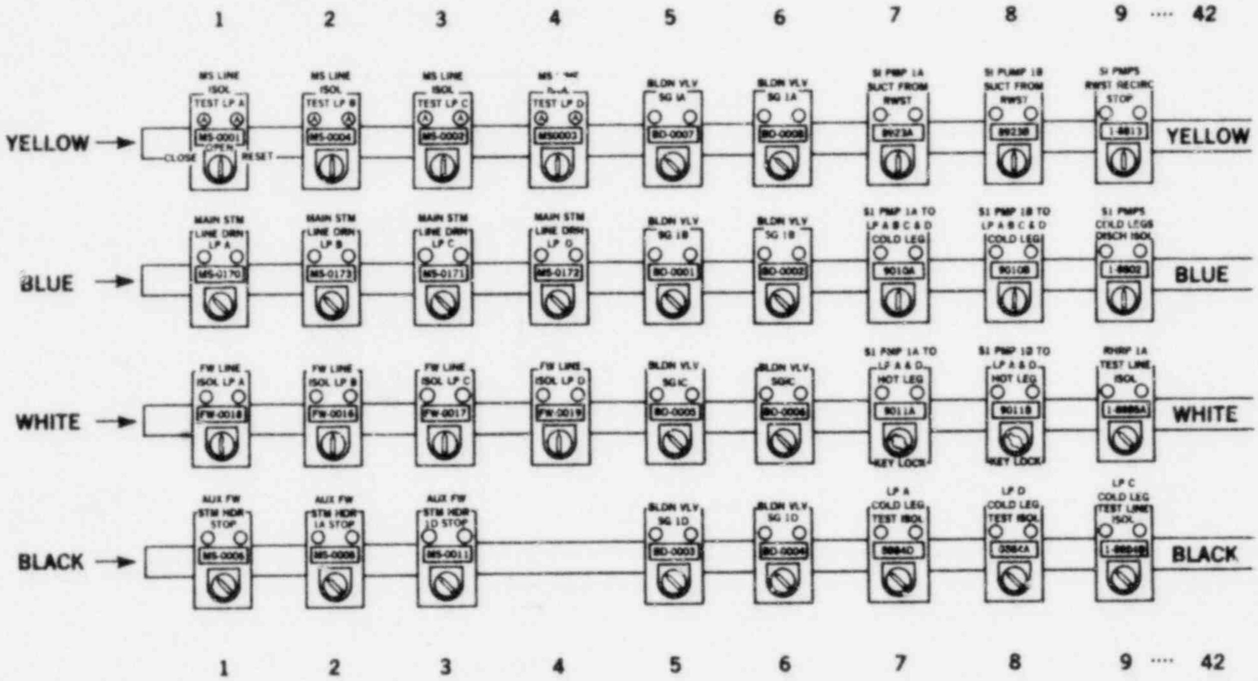


Figure 3-5. Part of the Lower Safeguards (ESF) Panel at a PWR

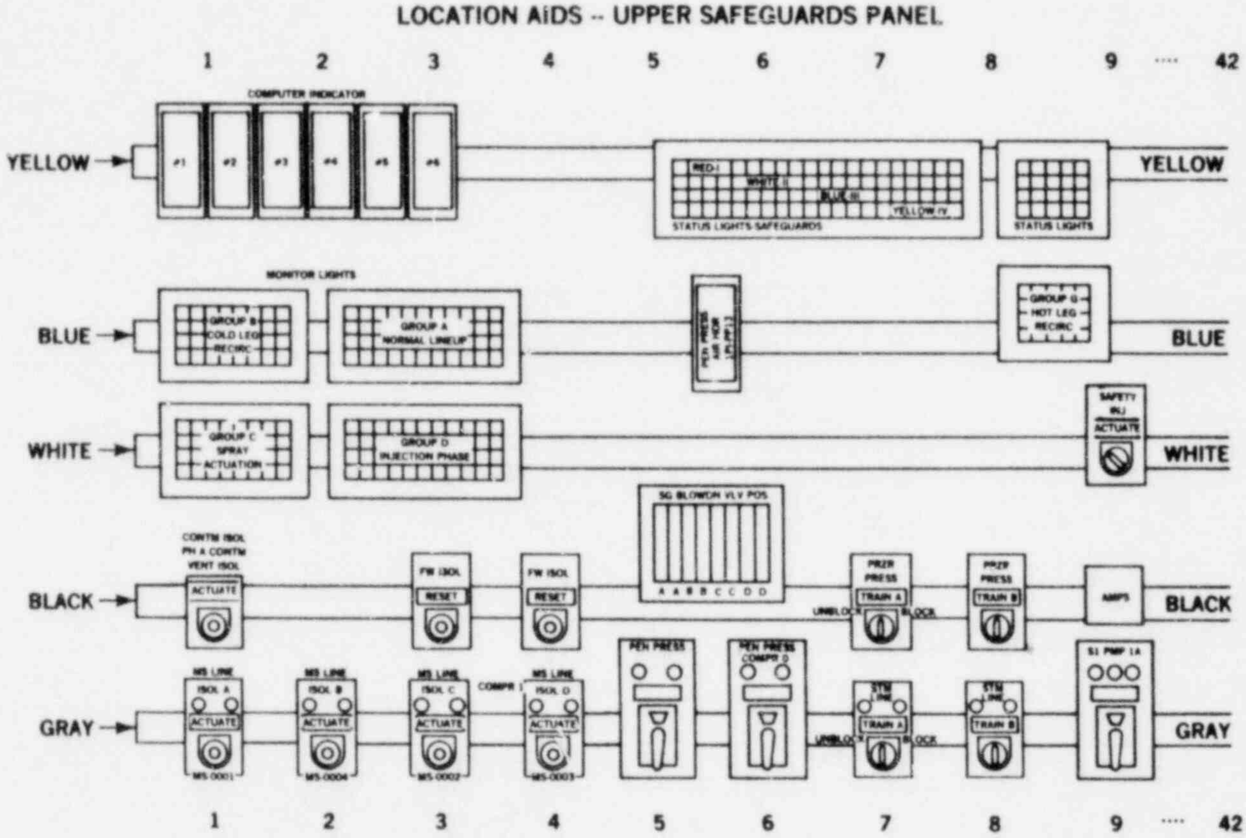


LOCATION AIDS -- LOWER SAFEGUARDS PANEL

Which is easier to find?

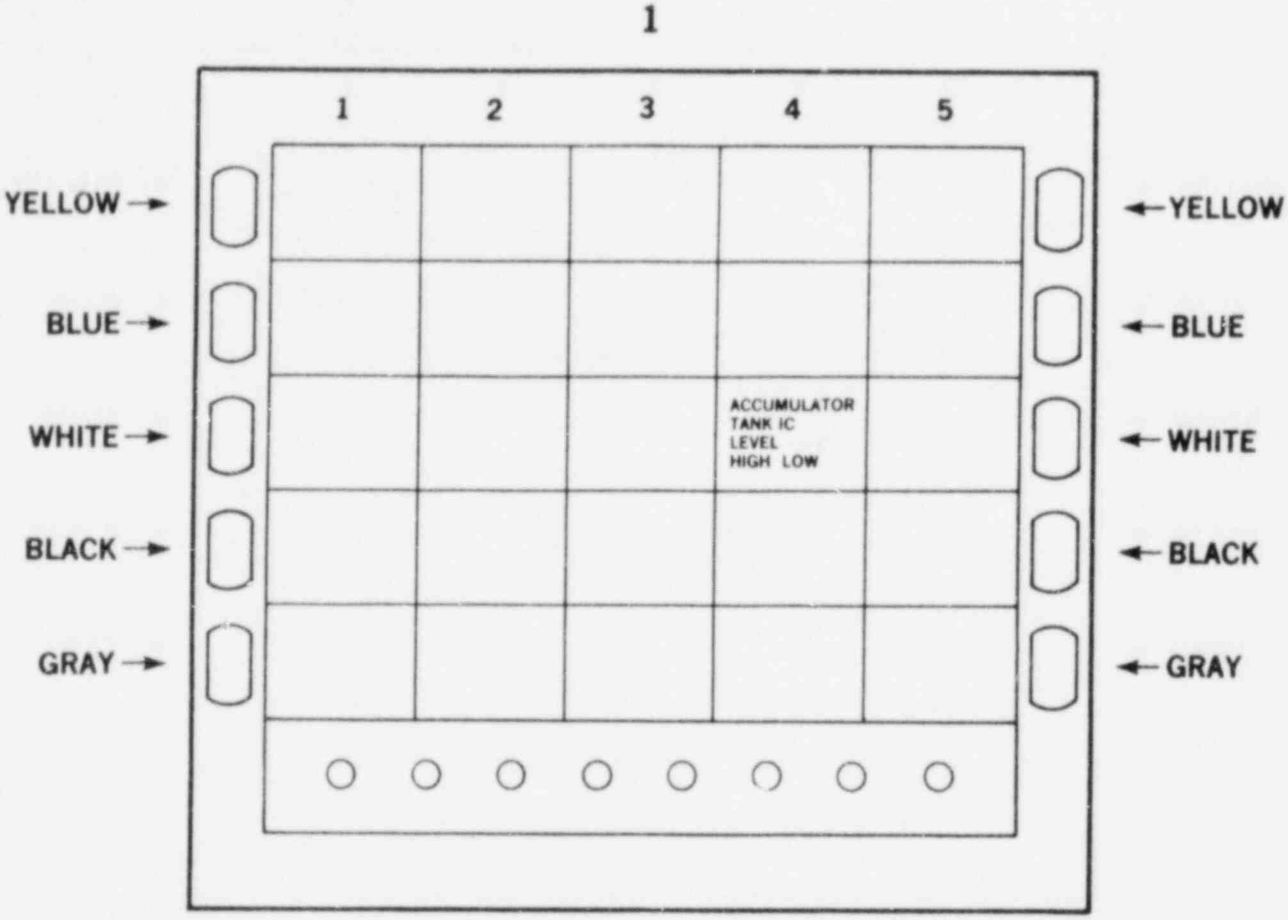
(1) Switch 8923B, or
(2) Switch 8923B (L-Yellow 8)

Figure 3-6. Location Aids for Lower Safeguards (ESF) Panel in a PWR



Problem: Quickly Find FW ISOL (U-Black-4)

Figure 3-7. Location Aids for Upper Safeguards Panel in a PWR



Problem: Quickly find ANN-1 White 4 when it may be one of several blinking annunciator indicators on several annunciator panels

Figure 3-8. Location Aids for Annunciator Panel #1 are the Federal Standard 595A (1974) identification numbers for the preferred shades of lusterless paint (used to avoid glare).

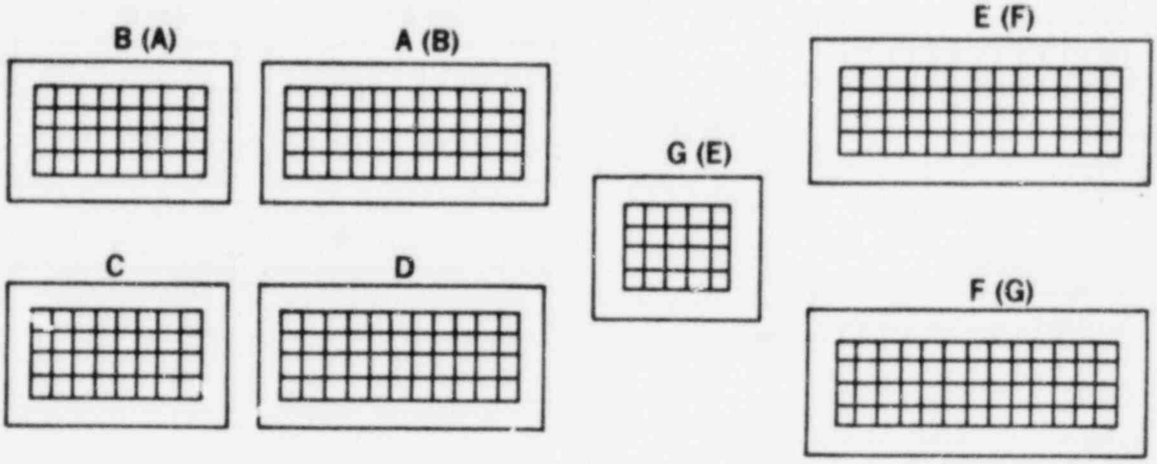
As an example of how the location aids would be used, assume that the operator had to locate Switch 8923B. Figure 3-5 shows how time-consuming such a search can be even with only one-fourth of the panel included in the drawing. For Figure 3-6 the written instructions would tell the operator to "turn on 8923B (L-Yellow 8)." (The L stands for the "Lower" panel.) One can readily see how this system would help in locating controls.

Although the personnel at the plant where this suggestion was first made thought this idea a good one, we were told it was not implemented because of management's objections that colored strips would ruin the aesthetic quality of the panels. (Letters could be used, rather than colored strips, with some reduction in location efficiency.)

In some plants, commonly accepted labeling conventions are violated. For example, the schematic in Figure 3-9 shows that the labeling of the six monitor panels on the ESF board in one PWR is not in accordance with the reading habits of the users: most of us read left to right and top to bottom. The figure shows the existing labeling, with the suggested labeling in parentheses. Correction of this type of poor labeling would involve minimal expense, but has not been done.

In another plant, on one panel there are four different labels used for "Pump." They are PU, PP, PMP, and PUMP. This may seem minor, but when one multiplies this one source of confusion by many, the information handling load on the operators increases materially.

Strong populational stereotypes are sometimes violated in the design of labels for displays. For example, in one PWR there are side-by-side



Monitor Panels: Current and (Suggested) Labeling

Figure 3-9. Violation of Reading Stereotype in Labeling of Monitor Panels displays showing Refueling Water Storage Tank (RWST) level and containment sump level (Figure 3-10). The indication for RWST level agrees with our populational stereotype: as the water level falls in the RWST, the pointer drops and shows a decreasing level. However, for the containment sump level display, the lowest level of water in the sump (4 inches) is shown by the top indicator and the highest level (205 inches) by the bottom indicator. This indication is the reverse of what one would expect.

An extreme example of mislabeling occurs in one plant in the labeling of switches for the primary coolant loops on the reactor control board (Figure 3-11). Switches for controlling valves associated with primary coolant loop A are logically labeled A, but switches for loop B are marked D, those for loop C are marked B, and those for D are marked C. The operators have attempted to cope with this by memorizing a mnemonic aid: All Dogs Bite Cats.

The major effects of inadequate labeling are increases in the perceptual and interpretational demands of the job, with the result that more errors occur and longer times are required to do the right things. This is one area in which substantial improvement in human factors could be achieved in existing plants with relatively little expense.

Inadequate Indications of Plant Status

Often the information conveyed by indicators of plant status requires excessive interpretation by the operators. At one plant a monitor panel was designed with the excellent intent that all the indicators would be illuminated when the monitored functions were in their proper modes so that the operator could quickly detect a deviant state of any individual

RWST - Refueling Water Storage Tank

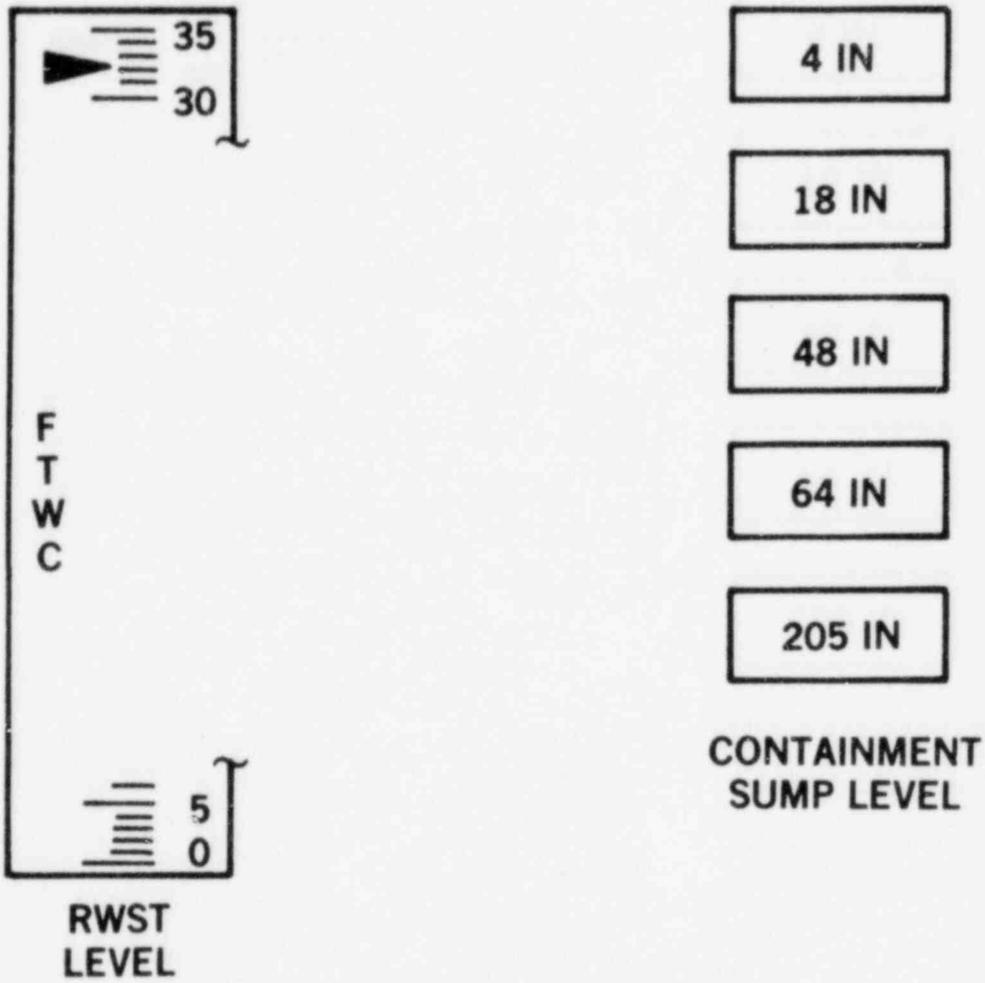
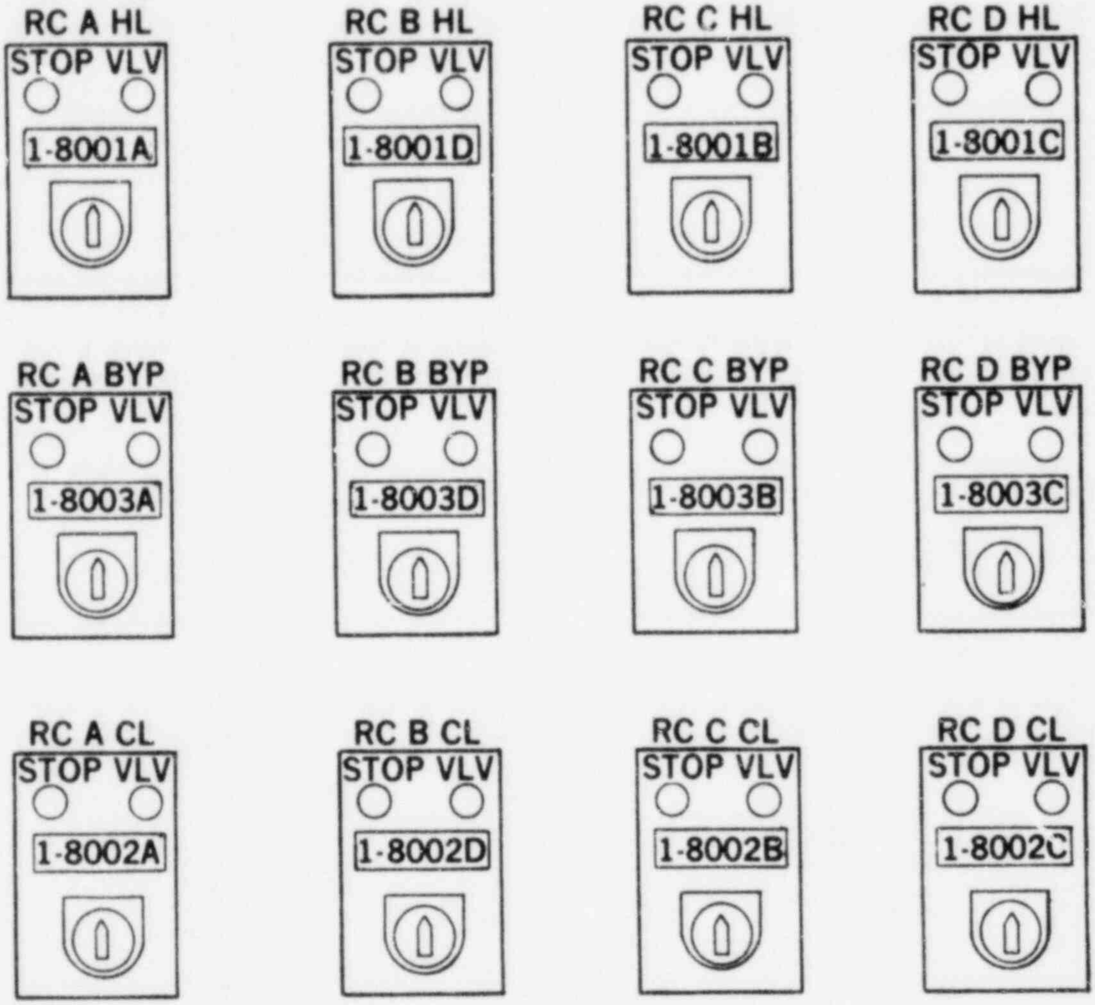


Figure 3-10. Two Adjacent Displays with Conflicting Representations of the Real World



(Green light is at left; red light at right.)

Figure 3-11. Reactor Control Switches for the 4 Primary Coolant Loops function. However, changes were made, and the operators had to learn that for condition so and so, all the indicators in Panel 1 should be illuminated--with a few exceptions. These exceptions would change for different conditions. The memorization task for the operators proved impossible. Interviews with these operators established that they temporarily memorize all the indications to pass examinations and then forget them.

The 1979 TMI accident illustrates several problems related to the requirement for excessive operator interpretation (Kemeny, 1979). One such problem is related to the closing of the pressure-operated relief valve on the pressurizer. The problem occurred because an amber light in the control room comes on when an automatic signal is sent to a solenoid to cause it to close the valve. The light does not monitor the position of the valve. The operator is supposed to remember that the amber light may not mean that the valve has closed (Rubinstein, 1979). Another problem occurred because there was no direct indication of flow in the auxiliary feedwater system. "The indicator the operator was supposed to watch to see if there was auxfeed was a pressure gauge on the discharge header. This is misleading if the auxfeed valve is closed because the pressure would be high and yet the flow is zero" (Sugarman, 1979). Again, too much interpretation is required for high levels of human reliability.

Another problem is the color coding of status lamps which follows the convention established in fossil fuel plants rather than ergonomics guidelines. In connection with NPP valves and switches, the color red routinely indicates the open status of a valve or the closed position of a swtich; i.e., in both cases red indicates flow. Green indicates the closed position of a valve and the open position of a switch; i.e., no flow. The use of these colors is contrary to our populational stereotype of green for go and red for no-go, and does not permit a more useful employment of these colors, as described later.

Other colors are also used in NPPs, but without uniformity. In some cases yellow is used for cautionary indications; sometimes blue is used for caution. Sometimes the same color is used for different indications within the same plant. At one plant a blue lamp was substituted for a red lamp to indicate an open valve. Added to these problems is the poor maintenance of color quality at some plants where formerly red or green filters had become so faded that the indicators looked white, a color that has an entirely different meaning. In one case we observed a formerly red lamp that now appeared dull green, a serious ELS.

These examples show that there are problems even with the simple use of color for status indications. More central to the operator's problem is that the typical NPP status lamp does not indicate the normal status of the monitored item for the steady-state operating mode of the plant; i.e., when the plant is in the normal power generating condition. This lack of information materially reduces the probability that an operator will detect unannunciated deviations from normal operating conditions.

A very practical method of presenting both the actual and the normal status of a control is used on shipboard reactors. The status lamps are shape-coded as well as color-coded. Two shapes are used: bars and circles. A pair of horizontal bars indicates that a valve is closed or a switch is open (the no-flow position). A circle indicates that a valve is open or a switch is closed (the flow position). These indicators are backlighted in either green or red, with green indicating the normal state and red the nonnormal state. Thus, a green illuminated circle indicates that a valve is open, or a switch is closed, and that this is the normal state for that item. If the circle were backlighted in red, it would indicate that the valve was open, or the switch closed, and that this was not the normal state for that item. In most cases a red indication would be the equivalent of presently used tags.

This shipboard philsophy is used at one BWR we have visited. When the reactor is in the normal operating power mode, a red indication means that the status of the indicator is not normal for that condition of the reactor, and a green indicator means that the status is normal. The bars and circles give the actual status. With this system, the operators know that the red and green cues pertain only to the normal operating power mode; for other conditions of the reactor, they have to depend on the bars and circles only and ignore the color cues. With new plants, even this limitation could readily be overcome with computer technology. It would be possible for a computer-based system to sense the plant's state, along with that of any given valve, switch, etc, and to indicate via the red or green light whether the state of the component is the appropriate state for the present operating conditions of the plant. The superiority in terms of human reliability of this design over the standard type of display (red means open and green means closed) is obvious. In a typical reactor, there may be several hundred dual-lamp indicators for status of MOVs, pump motors, and other equipment. Some of this equipment should normally be off or closed, and some should normally be on or open. The probability that an operator will detect nonnormal status is extremely low unless some other indication has directed the operator's attention to a specific pair of lamps.

In military and NASA systems, a green board philosophy is used where possible. In such a system, for the normal operating condition all indications for a given subsystem or function must be "green for go." Thus, any nongreen indication stands out and is much more likely to be detected. As far as we know, this excellent practice has not been followed in any commercial power plant. Such a practice, combined with functional groupings of displays, would materially improve the probability of detection of unannunciated deviant indications.

Apart from the above problems in color coding, significant numbers of people have various degrees of color discrimination weakness. The most common form of so-called color-blindness is red-green confusion experienced by about 7% of American males but by less than 1% of American females (Hilgard et al, 1971, p 119). This is a special problem since red and green are such central colors in the color coding schemes in most NPPs. However, this problem is recognized and utilities have color-blindness tests to screen for persons with such problems. For other process industries which may not employ such tests, forms of coding not relying on color discrimination would be useful.

An additional problem in status indications is the lag in display feedback to an operator performing rod control or chemical concentration adjustments. Because of the lag in system response, the operator must anticipate the response on the basis of his familiarity with it. While highly skilled personnel can perform this kind of operation reliably most of the time, errors do occur. Instrumentation techniques are available (Frost, 1972) that can predict system response so that the operator can tell what the terminal system state will be when he neutralizes his controls. The requirement for the operator to compensate for system lags is thus eliminated.

Presentation of Nonessential Information

The presentation of too much information, especially of excessive annunciated indications, is a problem in the operation of NPPs. There are hundreds of annunciator indicator panels in a typical control room, and the operators complain about the constant clamor of auditory alarms, most of which convey no real emergency message.

In discussions with design personnel, one gets the impression that each designer of a particular subsystem insists that his subsystem is so important that it must be annunciated. The result is an ineffective signal-to-noise ratio in the control room. Operators have been known to cancel both the auditory and blinking lights of an annunciated signal without even looking up at the annunciator lamp. Costly errors have occurred as a result. The reason for this behavior is as familiar and predictable as Aesop's fable about the boy who cried "wolf" once too often.

If the significance of each annunciated display in a plant were reviewed, many of them could probably be silenced as not requiring immediate action. Usually the blinking of the display (without an auditory alarm) is adequate to attract the operator's attention in time. Reducing the number of auditory alarms would lessen the tendency of operators to use the "kill switch" without attending to the display.

Inadequate Labeling and Status Indications of Manual Valves

Poor labeling can lead to errors in any aspect of NPP operations, but the most serious consequences are often associated with valving operations. There have been cases in which an operator closed a valve in the wrong reactor system because the labeling did not clearly indicate the relationship of the valves to the reactors. Errors such as this occur because the labeling may be difficult to read, located in an inconvenient spot, or missing altogether.

In addition to the problems of inadequate labeling, a source of uncertainty in valving operations is the lack of indication regarding the normal position of the valve, that is, its correct position for the steady-state operating mode of the plant. The operator may have to refer to some separate information source to ascertain whether the valve is in the correct position, a requirement that could lead to reversal or other errors.

If each valve had some indication of its normal operating position, (as defined above), errors in locating the appropriate valve and in deciding if the valve were in the proper position would be materially reduced. There are several ways of indicating normal operating position. For example, in the case of rising stem valves one could place on each valve a sketch of the valve stem in its normal position so that the operator could see if the valve stem matched the sketch. This inexpensive feature would materially reduce errors of commission.

At one plant we visited, a clever, inexpensive coding scheme used standard NPP meanings for the colors red and green. Green plastic tags were fastened to the handles of manual valves that were normally closed. Red plastic tags were used on manual valves that were normally open. This coding scheme eliminates the requirement for long-term memory of what the normal operating status is for each such valve. These plastic tags are permanently attached, and clearly different from the temporarily attached yellow cardboard tags which signify valves taken out of service.

In a plant which uses a color code of green for normal and red for nonnormal, a set of four tags for each valve designed to mimic the cue representations in the control room could be implemented. These tags could have green or red circles and green or red bars signifying normally open, nonnormally open, normally closed, and nonnormally closed valve states.

Job and Task Instructions

The PSF, "Job and Task Instructions," includes written or nonwritten procedures, written or oral communications, cautions and warnings, work methods, and plant policies (sometimes called shop practices). Although all of these are important for reliable human performance, our comments here are directed mainly towards written procedures, work methods, and plant policies. Oral communications are described in later chapters.

One of the most important work methods is the correct use of written procedures and checklists. If any task is performed without step-by-step reference to written procedures, errors of omission are much more likely. Also, if a checklist is used improperly, as when someone first inspects several items of equipment for proper status and then checks the checklist items all at once, errors of omission are again very likely. Chapter 15, "Recovery Factors and Administrative Control," provides quantitative estimates of the effects of these improper work methods.

Earlier, in discussing "Organizational Structure and Actions by Others," we stressed the importance of a proper and enforced administrative control system for the restoration of equipment after maintenance. This falls under the PSF of situational characteristics. In our opinion, the most important plant policy related to reliable restoration activities is the requirement to tag all valves removed from normal operating status, and to remove all tags when the valves are restored. The implementation of this policy involves a system of record-keeping in most NPPs which is inadequate in view of the importance of proper restoration of equipment.

As described in Chapter 15 (p 15-10), there are different levels of implementation of tagging policy. In terms of manpower expenditure and other expenses, the cost of the best level of tagging is small when compared with the losses prevented. A plant that uses a rigidly controlled tag inventory is less likely to experience unavailabilities because of failures to restore valves to their normal positions after maintenance. Effective plant discipline, enforced by supervision and management, is necessary for proper functioning of this best level of tagging procedure.

The performance models in Part III of the handbook show that estimated probabilities of unrecovered human errors can differ by factors of 100 or greater depending upon the type of written procedures and the related work methods used. Experience in many industries shows that well-written procedures and good work methods can often compensate for less-than-adequate human engineering of equipment. However, as stated in WASH-1400 (p III-64), "The written instructions [in NPPs] do not conform to established principles of good writing; they are more typical of military maintenance procedures of approximately 20 years ago."

Written procedures that are difficult to read, difficult to locate, or inconvenient to use are seldom used. At some plants, emergency procedures are not easily distinguishable from the other procedures; and, once located, a <u>specific</u> emergency procedure is difficult to find because there are no tabs or other indexing methods to assist the operator. Finally, the format and content of the <u>typical NPP procedures</u> are not conducive to easy use.

The reading problem has been studied intensively by the National Institute of Education, which reports that a substantial proportion of the U.S. population does not read well enough to function in society. Studies sponsored by that institute report that, "Some 12 million people 14 years of age and older cannot read as well as the average fourth-grader, yet seventh-grade reading ability is required to perform such skilled or semi-skilled jobs as machinist or cook" (Eaton, 1974). These and other data support our contention that NPP procedures should be written so that they are easier to understand and use. To reduce errors in using written procedures, the writing style should place minimal demands on reading ability. We estimate that the writing style of the typical NPP procedures requires about a Grade 12 reading level. We suggest that a Grade 4 reading level would be useful for maximum reading comprehension.* This is particularly important for highly stressful conditions, during which difficulty in comprehension can result in disorganized behavior as well as in specific errors.

It is not difficult to improve the readability of written procedures, even to write to the Grade 4 reading level. For example, Pyrczak and Roth (1976) studied the readability of directions on nonprescription drugs and found that the required reading level was usually Grade 11 or higher. They showed how some relatively simple rewriting could improve the readability to the Grade 4 level. The following statement on the bottle of a nonprescription drug is the Grade 11 or 12 reading level: "WARNING: Keep this and all medicines out of children's reach. In case of accidental overdose, consult a physician immediately." Their Grade 4 version

---

*This handbook is written at the college level, as it conveys concepts and principles, instead of simple instructions.

of this warning would be: "WARNING: Keep this and all medicines out of children's reach. If someone takes too much by accident, talk to a doctor right away." This example shows that writing at the Grade 4 level does not have to be insulting to the reader, a misconception by some who do not understand that the intent is to communicate reliably even under abnormal conditions.

Technical written procedures can be improved easily. For example, in one unpublished study by a Department of Energy contractor, the procedures for nuclear weapons assembly were revised so that the required reading grade level was reduced from Grade 11 to Grade 8. This was done by using photographs and drawings, by reducing the number of words by about one-half, and by reducing the average number of words per sentence from 16 to 8.5. Although the new procedures were about 50% costlier than the standard procedures, this increase was far outweighed by the reduction in human-initiated defects.

It has been known for some years that a columnar type of format for technical instructions is superior to a narrative format. In a 1969 study by Haney, experienced technicians familiar with the usual narrative type of format were tested with a columnar type of format (after a short practice session). They made one-third fewer errors with the latter format even though it was unfamiliar. With this study in mind, we rewrote part of the emergency LOCA procedures of one plant in a columnar format (Swain, 1975). Figure 3-12 shows part of the procedures for immediate manual actions. When the example was shown to the operating personnel and supervisors at one NPP, they liked the idea and adopted it. This format is discussed further in Chapter 14.

5. IMMEDIATE MANUAL ACTIONS

Step	Check	Indication or Item Manipulated	Activity	Result/Feedback
1	○	SI Initiation	Verify	SAFETY INJECTION ACTIVATED (Ann-7 Yellow 3)
2	○	REACTOR COOLANT PUMP 1A, 1B, 1C, 1D (U-Grey 84-87)	TRIP 4 Switches	TURBINE TRIP REACTOR TRIP (Ann-6 Grey 1)
3	○	GROUP A MONITOR LIGHTS (U-Blue-3)	Verify	Dark except: (1) N2-ACC 1880 (Yellow 4) (2) RHR HX2 1-0807 (White 3)

NOTE: This format is based on Swain (1975). Some revisions have been made for simplification.

Figure 3-12. Steps from Columnar Style Format for NPP Written Procedures

There are several advantages to the columnar format. First, many words can be omitted (conjunctions, articles, etc), resulting in a substantial gain in the signal-to-noise ratio. Second, important information is easy to find -- not buried in a sentence or paragraph where it might be overlooked. Third, the columnar format forces the writer to concentrate on what indications are presented to the user, what decisions he has to make, and what control actions he has to take. Fourth, with provision for checking off each item as completed, errors of omission are much less likely. Fifth, since such procedures are more convenient to use, it is more likely that they will be used.

Apart from problems in content and format, one of the most serious problems with NPP emergency procedures is that often there are too many instructions that are not safety-relevant. Much of this safety-irrelevant information concerns the reduction of monetary loss. We observed a talk-through of emergency procedures at one plant by a highly skilled and experienced shift supervisor. He performed under ideal conditions--no real stress present and no decision-making required. Yet he just barely managed to get through the procedures on a timely basis despite his exceptionally high skill level; there were too many tasks to perform within the allowed time. The money-saving instructions could have been put in a later section of the procedures, so that all of his initial effort could be devoted to the really critical issue -- the safety of the plant.

Stressors

Stress can be psychological, physiological, or both. Sometimes it is not possible to differentiate between the two. We have defined a stressor as "any external or internal force that causes bodily or mental tension." This definition allows an optimum level of stress as well as nonoptimum levels. This is a more general definition than the one given by Welford (1974) who states, "stress appears to arise whenever there is a departure from optimum conditions which the organism is unable, or not easily able, to correct."

Our reaction to a stressor is the stress we feel. Stress per se is not undesirable. As we will show later, unless there is some stress, nothing is likely to be accomplished in a work situation. Through common usage, the word "stress" has acquired a negative connotation because we tend to think of situations with high, incapacitating levels of stress. This is the kind of stress we wish to avoid in NPP operations, whether the stress is psychological or physiological.

Psychological Stress

Table 3-3 lists some psychological stressors. Some of these are clearly undesirable, but many are acceptable or even desirable in some limited amount.

Depending upon the level, psychological stress can be either disruptive or facilitative. Disruptive stress is the result of any stressor that threatens us, frightens us, worries us, angers us, or makes us uncertain, so that usually we do worse. The qualifier "usually" is necessary because of the large differences among individuals in response to these stressors; even the same individual reacts differently to the same stressor at different times.

TABLE 3-3.* Psychological Stressors

Suddenness of onset

Duration of stress

Task speed

Task load

High jeopardy risk

Threats (of failure, loss of job)

Monotonous, degrading, or meaningless work

Long, uneventful vigilance periods

Conflicts of motives about job performance

Reinforcement absent or negative

Sensory deprivation

Distractions (noise, glare, movement, flicker, color)

Inconsistent cueing

*From Table 3-1

We can use the word "arousal" for facilitative stress -- the result of any stressor that alerts us, prods us to action, thrills us, or makes us eager, but not too much. Again, a qualifer, "not too much," is necessary; if the stressor becomes too strong, it can have a disruptive effect. As with the response to disruptive stressors, there are large individual differences in what is felt as facilitative stress. A work situation that provides sufficient arousal for some people is seen by others as dull and montonous. Other things being equal, the higher the levels of education and technical skills a person brings to a job, the more arousal he requires. There is no "exciting and challenging" work per se, and there is no "dull and unchallenging" work per se; these are the judgments of people who differ in their perceptions.

Dealing with stress, or even getting people to agree on what stress is, is not easy. Figure 3-13 shows that when one plots stress level against performance effectiveness, the plot is not a linear one. With extremely high levels of stress (as exemplified by life-threatening emergencies), the performance of most people will deteriorate, especially if the onset of the stressor is sudden and the stressing situation persists for long periods (Berkun et al, 1962). Even when an escape route is obvious, some people will freeze up. A few people, like Audie Murphy, (the most decorated American soldier in World War II), will behave in an exemplary manner and do the right things at the right times. Regrettably, the Audie Murphy type of behavior is not universal under highly stressful situations (cf Berkun, 1964; and Ronan, 1953).

Figure 3-13 also indicates that at very low levels of stress, performance will not be optimum. There is not enough arousal to keep a

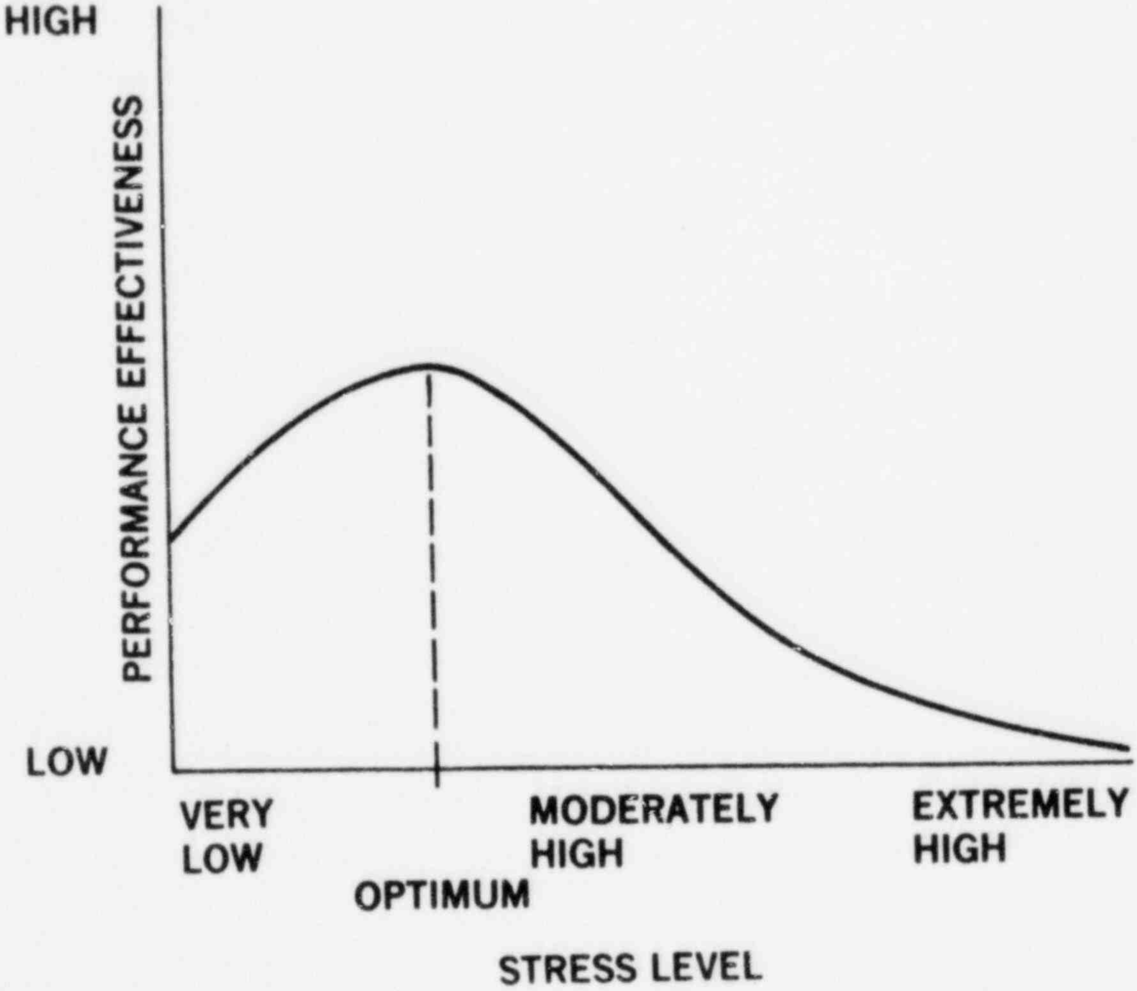


Figure 3-13. Hypothetical Relationship of Psychological Stress and Performance Effectiveness

person sufficiently alert to do a good job. Some people tend to drowse on the job, or their level of attention and job involvement is materially reduced.

The curve also shows that there is a level of stress at which performance is optimum. This optimum level of stress is difficult to define--it varies for different tasks and for different people. All we know is that the shape of the curve as shown is generally correct (cf Appley and Trumbull, 1967). This means that the tasks assigned to NPP personnel should be neither too boring nor so demanding that serious human errors are inevitable. With good ergonomics in the design of the plant and with properly skilled and practiced personnel, one has the best chance of avoiding both ends of the stress curve.

One of the difficulties in generating the stress models in Chapter 17 is that it is difficult to know just where on the stress curve certain unusual events in NPPs will fit. In our model, we assume that a design-basis LOCA* is near the far right of the curve. Such a large LOCA would meet several requirements for classification as a severe stressor. First, based on their training, operators recognize the possible consequences to the public and to the plant of an uncontrolled large LOCA. Second, it would be totally unexpected. Interviews with plant personnel and with NRC personnel indicate that no one thinks a design-basis LOCA will ever occur. If this well-founded opinion proved false and a large LOCA did occur, the most likely reaction of the operators would be one of sheer disbelief. We call this the incredulity response. It has been

---

*A design-basis LOCA is one in which one or more large coolant pipes suddenly experiences a guillotine type of break. It is also generally assumed that this break occurs when the emergency coolant inventory is at the lowest operating level allowed by NRC technical specifications.

observed in other work situations. For example, in one refinery the first indication the control room operator had of a serious fire was that many alarms occurred and many instruments behaved abnormally. This operator's first response was to run upstairs and demand of an instrumentation technician, "What's wrong with my instruments?" By the time he returned to the control room it was too late to take action that might have reduced the loss due to the fire.

Finally, operators rarely practice responses to simulated large LOCAs (or other unusual events) after their formal training. Yet the only way to minimize the incredulity response is to provide frequent drills so that the operator will be well practiced in responding to low-probability events. Unfortunately, this internal PSF, "State of Current Practice or Skill," is not at an optimum level for most NPP personnel.

We judge that unusual events that are less threatening to NPP personnel than a large LOCA should be placed around the moderately high stress part of the curve in Figure 3-13. Examples might include certain hot shutdown activities and other tasks that place time-stress on a person, but without disrupting factors such as fear, anger, and uncertainty.

There are two important problems for human reliability under high levels of stress -- (1) man tends to revert to his populational stereotype, and (2) he tends to perseverate among a very few response alternatives. When we are in an emergency and are experiencing stress, as evidenced by tightening of stomach and sphincter muscles, pounding of the heart, dryness of the mouth, etc, we tend to do "what comes naturally" or revert to our populational stereotype. This means that we will see things as we customarily see them and will respond in the way we are accustomed to responding. If some man-machine interface violates these ingrained habits of perception and response (e.g., an emergency switch that must be flipped up for "off," or a manual valve that must be turned counterclockwise for "off"), the probability of inappropriate action is extremely high. We estimate that, under highly stressful conditions, and even despite extensive training, the probability of human error in such cases ranges from .1 to .9 if the equipment violates a person's populational stereotypes.

Different populations have different populational stereotypes; for them, designs that conform to U.S. populational stereotypes can result in errors. For example, in European countries the convention is for toggle switches to be flipped up for "off" and down for "on." In other countries, the natural order of things corresponds to their reading stereotypes, which can be quite different from those in the U.S.

Whereas the problem of populational stereotyping can be solved by appropriate human factors engineering, the problem of response perseveration can be solved only by a combination of good design, training, and practice. Response perseveration is the term for the tendency to make some response (or a very limited number of responses) that is incorrect repeatedly. This may be in response to some unusual but not especially stressful event, as when a motorist (even an engineer!) repeatedly pumps the gas pedal when trying to start a car with a flooded carburetor.

Perseverative behavior has been observed in people under the severe stress of combat (Grinker and Spiegel, 1963), under realistic experimental conditions (Berkun et al, 1962, p 27), under the much less stressful condition of trying to troubleshoot electronic equipment under time pressure (Bryan et al, 1956; and Bond, 1970), and under other conditions in which the correct path of behavior is not clearcut. Ambiguity resulting in response perseveration can arise from inadequate presentation of information (a design problem), from lack of skills to process adequate information (a training problem), or from inability to recall and use the appropriate skills because of lack of continuing practice (also a training problem).

The low end of the stress curve (see Figure 3-13) has important implications for monitoring tasks. If a control room operator is not sufficiently aroused, he is less likely to detect deviations from normal before they result in some annunciated indications. If an operator's first indication of something untoward is an annunciated signal, he may not always be able to correct the situation on a timely basis (Seminara et al, 1976). This is in part a design problem, but it is also a problem of ineffective monitoring that develops when the operator is not experiencing enough signals to maintain arousal or alertness. This loss of alertness is called the vigilance effect (Figure 3-14). This phenomenon was noted in World War II by the British, who discovered that the maximum time a shipboard lookout could be kept on duty effectively was about one-half hour. After that, the probability of his detecting an enemy submarine or aircraft was unacceptably low even though his own life and those of his shipmates were at stake. Later research verified the vigilance effect and found that it applied also to some industrial inspection tasks in which the involvement of the inspector was essentially passive, such as in looking for defects when the number of actual defects was very low (one or fewer defects per 100 items) (Harris and Chaney, 1967, 1969; McCornack, 1961; and Fox, 1975).

In WASH-1400 we stated that the level of activity in a control room was usually such that the vigilance effect, or the low end of the stress

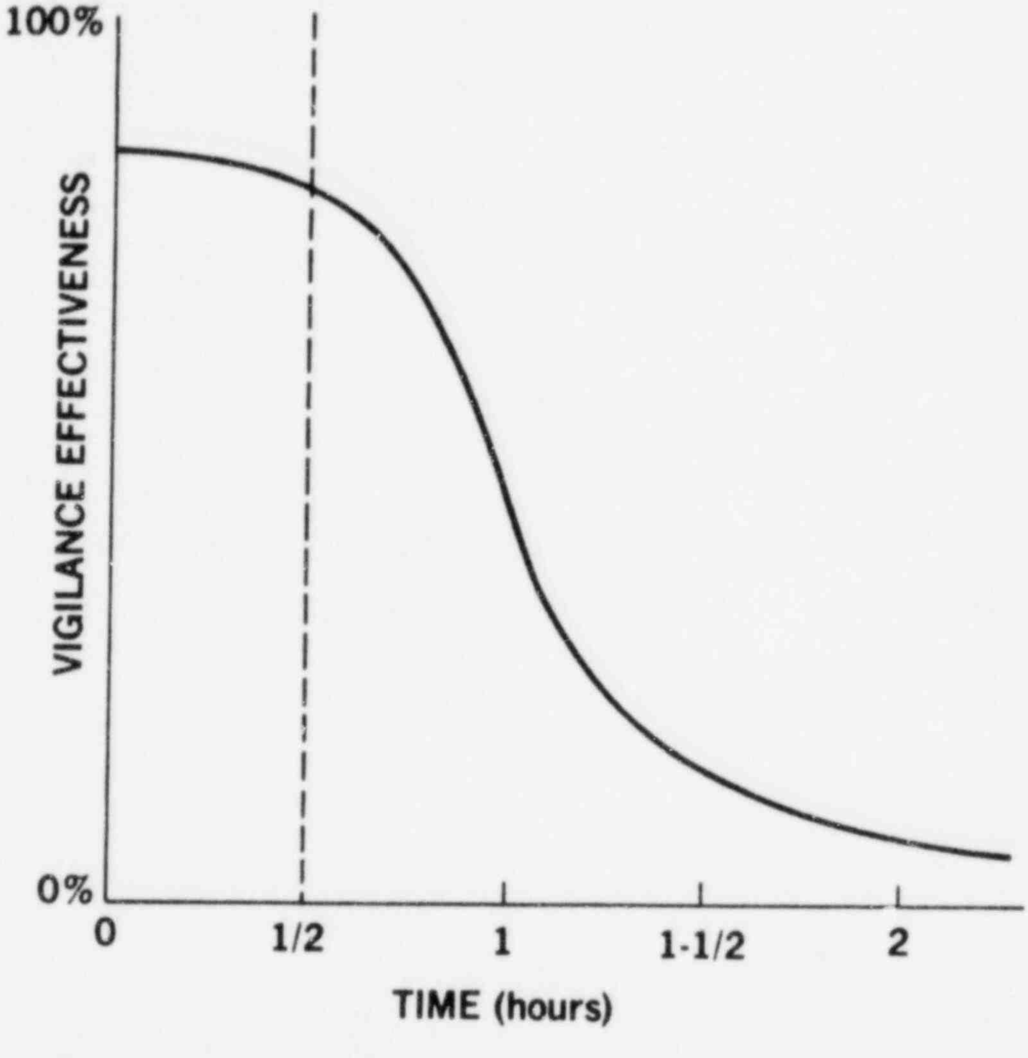


Figure 3-14. Vigilance Effect for Passive Tasks with Low Signal Rate

curve, did not apply. However, subsequent information we have gathered in observations of U.S. and European NPPs and from the EPRI Review (Seminara et al, 1976, p 18-6 to 18-9) indicates that, at least in some plants and especially during the evening and night shifts, operators often consider their work to be very dull, monotonous, and unchallenging. At one European plant there was even a request to install a television set so that the operators could watch TV programs when not otherwise busy. (The request was turned down.) Our modeling considers the effects of nonarousal in the low estimates of probabilities of detection of unannunciated deviant indications.

In summary, the effect of psychological stress in NPPs is a serious problem. It can be addressed effectively through a combination of sound equipment design, frequent practice, and responsible supervision.

Physiological Stress

Table 3-4 lists some physiological stressors. As stated, all of these stressors would be disruptive. We have already addressed the effects of fatigue (pp 3-9 to 3-12). The special problem of working in a radiation environment is discussed in Chapter 17, "Stress."

Few of the other stressors constitute serious problems in NPP operations. However, discomfort can be a highly disruptive PSF for certain maintenance tasks in which awkward positions must be assumed for access to components. Errors, especially errors of omission, can be expected to increase, particularly if such discomfort is combined with temperature extremes, as is sometimes the case.

Movement constriction and lack of physical exercise is a problem primarily in the control room. However, it is common practice for operators to walk around frequently not only to monitor displays but probably

Table 3-4.* Physiological Stressors

Duration of stress

Fatigue

Pain or discomfort

Hunger or thirst

Temperature extremes

Radiation

G-force extremes

Atmospheric pressure extremes

Oxygen insufficiency

Vibration

Movement constriction

Lack of physical exercise

---

*From Table 3-1

also just to get up and move around. Some designers of NPPs have misapplied this small problem of movement constriction and have argued that a standing operator is more effective than a seated operator. A sitdown console concept was changed to a standing one because of this mistaken belief. What the designers failed to consider were the PSFs of fatigue and discomfort. Furthermore, when operators need to sit, they will sit, even if this means sitting on consoles or other places where inadvertent manipulation of controls could result.

One final physiological topic is mentioned only because many people ask about it. This is the idea that one's "biorhythm" affects performance, and that each operator's biorhythm should be determined so that he is not assigned to critical or dangerous work on biorhythmically critical days. Extensive reviews of biorhythm theory (Wolcott et al, 1977; McConnell, 1978; and others) indicate that, while there are certain psychophysiological rhythms, the 23-day physical cycle, 28-day emotional or sensitivity cycle, and 33-day intellectual cycle postulated by this theory are not supported by any reliable evidence. However, there is evidence to suggest that the individual circadian cycles of operators do affect their performances (Folkard et al, 1979; Colquhoun, 1970; and Colquhoun et al 1968a, 1968b, 1969). Supervisory personnel in NPPs fail to take this factor into account in that they rotate shift changes on a weekly basis. Humans require from 4 days to a week to adjust to radical shift changes (ones that materially disrupt their established sleep schedules). The weekly shift rotation does not allow sufficient time for recovery. (See also Maurice, 1975, for the effects of rotating shifts on industrial errors.)

Summary of Human Reaction to Stress

When overburdened by a situation (task stress or speed stress), people respond to stress in any one or more of several ways as listed below (Edwards and Lees, 1973, p 20):

Queueing - delaying some responses during overload, with the intention of responding at a later time.

Omission - ignoring information or actions that are considered relatively unimportant.

Gross Discrimination - responding to gross aspects of signals and ignoring finer aspects; e.g., noting that the water level in the sump has risen but not noting the extent of the change.

Errors - processing information incorrectly.

Escape from Task - physical or mental withdrawal.

As can readily be seen, some of these responses are more detrimental than others in their consequences for a man-machine system.

Internal PSFs

Table 3-5 lists some of the internal factors of the individual in a man-machine system. Some of these PSFs are outside the control of supervision and management, but most are either the direct responsibility of the utility or can be positively influenced by utility policy.

In WASH-1400 (p III-64) we judged that the level of training of NPP personnel was outstanding. Based on our subsequent studies and on the EPRI Review (pp 18-9 to 18-14), it is apparent that this earlier judgment should be modified. We still believe that the training of NPP control room operators is good, but there is much room for improvement (Kemeny, 1979). Moreover, another EPRI report indicates that the training of

Table 3-5.* Individual (Organismic) Factors

Previous training/experience

State of current practice or skill

Personality and intelligence variables

Motivation and attitudes

Knowledge of required performance standards

Physical condition

Attitudes based on influence of family and other outside persons or agencies

Group identifications

*From Table 3-1

maintenance personnel is quite deficient (Seminara, Parsons, et al, 1979). As was the case in the training of military electronics personnel in the 1950s, NPP training courses include much theory that may not be necessary for plant personnel who perform operator, maintenance, or other hands-on activities. With limited amounts of time for training, and with costs between $100,000 and $200,000 to train each operator, the elimination of job-irrelevant training from the syllabus would allow more time for operationally oriented content. It is apparent from the EPRI reports that the training of NPP personnel needs a thorough reevaluation.

While there may be some reservations about the quality of training, there is definite concern about the PSF of "state of current practice or skill" of safety-related tasks. In WASH-1400 (p III-64) we were critical of the lack of practice provisions for safety-related tasks. Nothing has changed this view, and it is further supported by the EPRI studies. Interviews with operating personnel indicate that they get very little practice in coping with simulated emergencies. Their original training includes valuable practice in dynamic simulators, but once they are assigned to a utility, it is apparently assumed that what they learned in the simulator will remain with them forever. It is mistakenly believed by some that the required operator requalification every two years includes extensive practice of simulated emergencies in a dynamic simulator. However, there is sufficient leeway in the interpretation and application of NRC regulations for operator recertification that, a least prior to the TMI accident, it may have been the exception for an operator to receive such practice every two years. This is analogous to training a pilot very thoroughly in coping with inflight emergencies initially and then assuming that he needs no periodic practice to maintain this initial proficiency. In the case of commercial pilots, however, recertification is required by international agreement every six months, and this recertification must include practice of simulated emergencies in dynamic simulators.

In Figure 3-15 we postulate the general shape of the curve for loss of ability to cope with emergencies in the absence of practice (the solid line) compared with the continuing improvement that takes place with periodic practice (the dotted line). The time intervals for periodic practice by NPP personnel would have to be determined empirically, and the ratio of the time periods spent in dynamic simulators to the time spent on other simulation would also have to be determined.

The other simulation could consist in large part of talk-throughs and walk-throughs of emergencies and other unusual events (see p 4-12 for a description of these techniques). As noted in WASH-1400 (p III-64), we made an informal test using talk-throughs and, "It was found that operators interviewed could explain in general terms what they should do in postulated emergency situations, but they did not always appear to be sure of the locations of switches and readings on displays relevant to manual backup actions required in the event of failure of automatic safeguards systems. ...the lack of ability to 'talk through' appropriate procedures without hesitation or indecision potentially indicates lack of a clear plan of action should such emergency situations occur. Based on the above findings, relatively high error rates were consequently assigned to operator actions required soon after the onset of a major emergency such as a large LOCA."

Our conservative estimates of operator ability to respond properly under highly stressful conditions could be modified upward if talk-

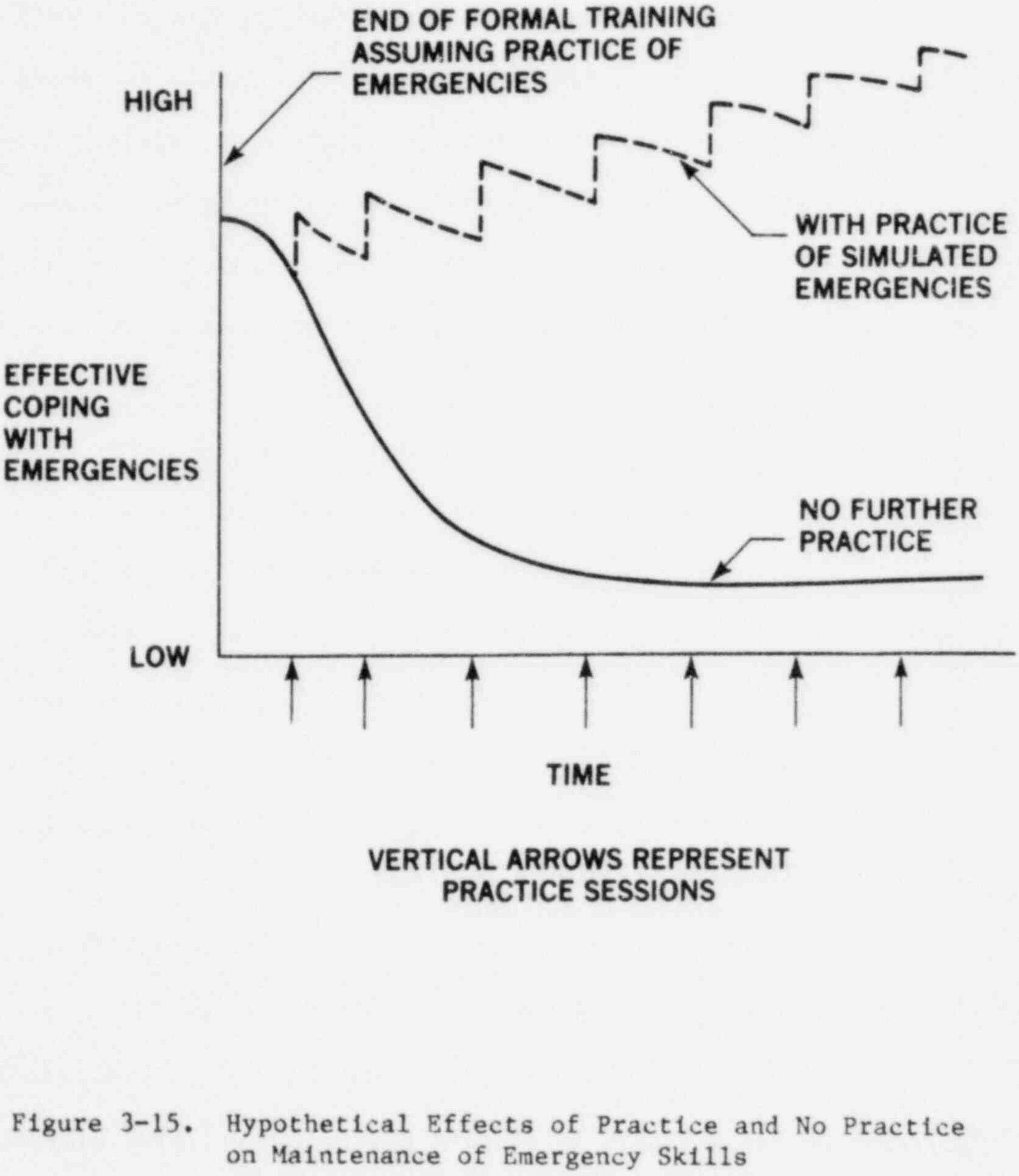


Figure 3-15. Hypothetical Effects of Practice and No Practice on Maintenance of Emergency Skills

throughs of these tasks were practiced frequently. To this end, we have urged the NRC to initiate formal tests using the talk-through method. Periodic testing by the onsite Inspection and Enforcement personnel would provide sufficient motivation for a plant to ensure that its personnel would practice coping with simulated emergencies.

Although personality and intelligence variables obviously influence human reliability in NPPs, these variables are not nearly so important as other PSFs, especially those related to practice of skills. However, there is a potential motivational problem related to selection and training content that is not fully job-related. If it is true that much of the nuclear theory currently given operator-trainees is really not necessary for reliable operator performance, two undesirable results occur. First, potentially competent persons are eliminated on the basis of job-irrelevant considerations. Second, those who successfully pass the training may be disappointed that the theory they have learned is not required, and they may feel that their skills and abilities are not being used. This could have a negative effect on their motivation.

The motivation and attitudes of the individual in an NPP obviously have considerable influence on how well he performs. From experience it is known that a well-human-engineered work situation plays an important role in operator acceptance of and enthusiasm for his work. Thus, application of sound human factors practices to NPP design and work operations would have a substantial beneficial effect on operator motivation and attitudes.

The last three PSFs from Table 3-5, the operator's physical condition, his attitudes based on outside influences, and his identification or affiliation with various groups, deal with influences that are not under the control of a utility, but are listed to show that the responsibility of a utility for the performance of its personnel does have limitations.

Effects of Good Human Factors Practices on Estimated HEPs

The estimated HEPs in this handbook reflect our evaluation of the effects of current design practices in NPPs on human performance. In our survey of NPPs before and during the WASH-1400 study, and in subsequent visits to other NPPs, it became apparent that no systematic consideration of human factors technology was incorporated in the design of man-machine interfaces, written procedures, or operational practices. Violations of conventional human factors practices (as outlined in MIL-STD-1472B, 1978) are the general rule rather than ocasional occurrences.

For several years, human factors experts at the American Institutes for Research, the Electric Power Research Institute, Human Performance Technologies, the Lockheed Missiles and Space Co., LUTAB, Risø National Laboratory, Sandia National Laboratories, and other institutions have pointed out that the design of man-machine interfaces in NPPs is not fully compatible with the capabilities, limitations, and needs of the personnel who function at those interfaces. There is no doubt that the incorporation of good human factors practices in the design of NPPs and related human operations and procedures could effect substantial improvements in human reliability.

Just how much improvement could be effected is obviously situation-specific. In Table 3-6 we have developed some conservative estimates of the benefits that would result from an across-the-board application of good human factors practices to NPPs. These estimated factors are not additive.

Table 3-6. Estimated Decreases in HEPs Resulting from Application of Good Ergonomics Practices to Nuclear Power Plants

If done:	Resulting Decrease in HEPs (Factors):
Good human engineering practices in design of controls and displays	2 to 10
Use of well-designed written procedures and checklists to replace typical narrative style procedures	3 to 10
Redesign of displays or controls that violate strong populational stereotypes	≥ 10
Redesign of valve labeling to indicate their functions (including a clear indication of the system with which a valve is associated) and also to show clearly their normal operating status	~ 5
Frequent practice of the appropriate responses to potential emergencies of other abnormal situations (practice includes periodic recertification in dynamic simulators and talk-throughs conducted at least once per month for the major potential problems)	2 to 10

PART II. METHOD FOR ANALYSIS AND QUANTIFICATION OF HUMAN PERFORMANCE

PART II. METHOD FOR ANALYSIS AND QUANTIFICATION OF HUMAN PERFORMANCE

The general method for the analysis and improvement of human performance consists of the following steps:

(1) Identify all the interactions of people with systems and components; i.e., the man-machine interfaces.

(2) Analyze these interfaces to see if the PSFs are adequate to support the tasks that people have to perform.

(3) Identify potential problem areas in equipment design, written procedures, plant policy and practice, people skills, and other factors likely to result in human error.

(4) Decide which problems have sufficient potential impact on the system to warrant changes.

(5) Develop candidate solutions for the problems.

(6) Evaluate the estimated consequences of these changes to ensure that they will improve system reliability and safety and that no additional serious problems will result from them.

This general method, which has been in use for some time in the human factors community, is called man-machine systems analysis (MMSA). The descriptive and analytical part is often called task analysis. The quantitative part uses a human reliability model to develop estimates of the effects of human performance on system criteria such as reliability and safety.

Chapter 4 describes the task analysis, which furnishes the raw data for the human reliability model described in Chapter 5. This model has been used in NPP reliability analyses in the U.S. and in Europe. A sample application of the model is presented in Chapter 5, and other applications are found throughout the handbook. An outline of the general methodology and some case studies are presented in Chapter 21. Chapter 6 discusses the use of HEPs to estimate component unavailabilities, and provides some examples of unavailability calculations. Other examples are found in Chapters 8 and 11.

The methods described in Chapters 4 and 5 are directed towards a complete MMSA in either its qualitative or quantitative mode. The qualitative use of MMSA is directed towards identifying error-likely situations in a man-machine system without attempting to assess the relative importance of any given ELS. When the latter assessment is performed, the MMSA becomes quantitative. Obviously, the qualitative use of MMSA is always a requirement for its quantitative use. Quantitative uses range from the incorporation of very simple scales (e.g., a 5-point ordinal scale of error-likeliness) to the use of HEPs described in the handbook.

The methods in Chapters 4 and 5 can also be used for analyses that are less than complete, as for example, when an analysis is done across all plants on some particular task or set of tasks to determine the interplant range of human reliability for that task. Chapter 21 presents some applications of these methods to arrive at approximate assessments and also presents applications for the following analyses:

Worst-Case Analysis - in which consistently high estimates of HEPs (e.g., the upper uncertainty bound rather than the nominal HEP for each HEP) are used to present an overly conservative assessment of the impact of human errors on a system.

Best-Case Analysis - in which consistently low estimates of HEPs (e.g., the lower uncertainty bound rather than the nominal HEP for each HEP) are used to present an overly optimistic assessment of the impact of human errors on a system.

Bounding Analysis - in which both the worst-case and best-case analyses are used to establish boundaries of the estimated influence of human performance in a system.

Chapter 4 (p 4-20) briefly describes a fourth type of analysis, sensitivity analysis, in which the estimated HEPs, dependence levels, or other indices of human performance are systematically varied to determine the effects of such variation on system outcomes. Chapter 7 and 21 include some examples of sensitivity analysis.

Although the above analyses are only intended to be useful approximations, they should be based on as detailed a task analysis as can be performed. To do otherwise risks overlooking important PSFs that could significantly affect human reliability. The most useful reliability assessments will be made by people who collectively represent high-level skills in human factors, reliability technology, and statistics. Unless all of these skills are represented, any assessment of the role of the human in a system may fail to identify and properly evaluate all of the human performance variables having potential impact on system safety or availability.

CHAPTER 4. MAN-MACHINE SYSTEMS ANALYSIS

This chapter describes the basic analytical methods used in the human factors community to identify and evaluate existing or potential human performance problems in man-machine systems. These methods furnish the raw material for a human reliability analysis, described in Chapter 5.

An adequate human reliability analysis is based on a thorough analysis of the operator's tasks in the context of the application; e.g., the nuclear power plant. Techniques for identifying ELSs and APSs in complex man-machine systems were developed in the early 1950s by Dr. Robert B. Miller and his associates at the American Institutes for Research. These techniques, collectively titled A Method for Man-Machine Task Analysis (Miller, 1953b), have been refined and expanded for application to human reliability analysis. The general term man-machine system analysis (MMSA) includes both qualitative and quantitative aspects of human performance assessment. The ten iterative steps in MMSA are listed in Table 4-1 and are discussed below.

Step 1 - Describe the System Goals and Functions

The purpose of this step is to see where people fit in with system goals and functions. What are people supposed to do to accomplish various system functions? Where are the points of interaction between the system and the people? In WASH-1400, these points were defined as the interfaces between equipment and people; e.g., manual valves, switches for motor-operated valves, displays to be read, provisions for calibrating setpoints.

It is especially important to understand the assumptions about people that are inherent in the design of each system. Usually these assumptions will not be stated, and must be inferred.

Table 4-1. Steps in Man-Machine Systems Analysis (MMSA)

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

The person performing an MMSA should not unquestioningly accept the definition and division of jobs and tasks as they exist in the plant, or in a plant being designed. For each system function, one must determine whether there is a reasonable division between those tasks that are accomplished by equipment, those by people, or those that are accomplished by an interaction of both. Too frequently, this allocation seems to have developed historically by trial and error rather than through systematic analyses.

Sources of information for this step include design requirements, planning documents, proposals, schematics, flow diagrams, written procedures, and interviews with system planners and people with experience in the operation of similar systems. For the evaluation of an existing plant, the information should be checked by visits to the plant.

Flow charts with a time baseline may be used to show the system functions for each major area in the plant. Flow charts can show how people fit into the system and what the design constraints are. (For preparation of flow charts, see Edwards and Lees, 1973 and 1974.)

Step 2 - Describe the Situational Characteristics

Situational characteristics of interest are those PSFs under which the tasks will be performed. Examples are air quality, general cleanliness, lighting, accessibility, union restrictions, and other performance shaping factors listed in Table 3-1 on p 3-5. Some PSFs may vary from job to job in a plant, but several will be essentially the same for several jobs. Sources of information include the documentation listed in Step 1, but the best sources will be interviews with management and supervisory personnel, observation, and interviews with people in the various plant work areas.

Step 3 - Describe the Characteristics of the Personnel

In this step the skills, experience, training, and motivation of the personnel who will operate, calibrate, and maintain the plant systems are identified. The capabilities and limitations of the people in a system must be understood so that they can be compared with the demands the system makes upon them. Any mismatch between these two sets of factors requires either a change in man-machine interfaces or modification of personnel characteristics through training and/or selection.

One important aspect of this step is to evaluate people's past experience with other systems in order to avoid transfer of habits that would interfere with reliable performance in the new system. At present there is no standardization of ergonomic considerations in NPPs. Possibilities for negative transfer of habits must therefore be evaluated when personnel are assigned who may have worked in other plants or trained on simulators where the PSFs differed materially from those in the plant in question.

For safety-related operations, it is important to evaluate the provisions for continued practice of responses to low-probability events such as a LOCA or an anticipated transient. Without practice, the readiness to handle such events will decrease, as explained in Chapter 3 (see Figure 3-15, p 3-69).

Step 4 - Describe the Jobs and Tasks that the Personnel Perform

Steps 4 and 5, jointly, constitute task analysis. Task analysis is an analytical process for determining the specific behaviors required of the human components in a man-machine system. The individual tasks, or steps in the tasks, become the limbs in the probability tree diagrams used for human reliability analysis.

Task analysis can be divided into description and analysis. This step deals with the descriptive part, and lists those PSFs related to (1) task and equipment characteristics and (2) job and task instructions. With the situational characteristics from Step 2, they describe the demands that each job places on the personnel.

There are many different formats for task analysis. The particular format used is unimportant; the important thing is to describe and analyze each task as necessary and to identify ELSs and APSs. Figure 4-1 shows a format (reduced in size) used in some early weapons studies that illustrates the kinds of factors to be considered in a detailed task analysis. Note that the format includes a descriptive part ("Task Behaviors") related to Step 4 of the MMSA and an analytical part ("Task Components") related to Step 5 of the MMSA.

There are five columns in the descriptive part of the format. In the first column, "Task or Step," one uses numbers to indicate the sequence of performance. Under the second column, "Instrument or Control," one lists each item that displays information to the operator or that must be manipulated. In the control room, for example, the annunciators, meters, chart recorders, and other items display information. The controls are mainly switches on the control panels and typewriter keys for the computer. For calibration tasks, the "Instrument" includes meters for measuring setpoints and the "Controls" include potentiometers for adjusting setpoints. Controls also include connectors and tools. In all cases, the labels on the equipment are used and are capitalized if that's the way they appear on the equipment.

In the third column, "Activity," one sets down action verbs describing the human actions to be carried out on the items in the second column. The action verbs should help identify the kind of display or control used. For

Job: _____ Task: _____ TASK BEHAVIORS (Description) Page ___ of ___ Pages TASK COMPONENTS

Subtask: _____ Conditions: _____ Task Analyst: __________ (Analytical)

Task or Step	Instrument* or Control†	Activity	Cue for initiation or completion of activity (immediate or delayed)	Remarks	Task or Step	Scanning, perceptual, anticipatory requirements	Recall req's (LT = long-term ST = short-term # = initiating cue absent or poor)	Interpreting req's	Manipulative Problems	Likely human errors (or other problem areas)	Check if confirmed

*Anything that displays information
†Anything that is manipulated

Figure 4-1. A Job-Task Analysis Format Used for Weapon System Studies (Swain, 1962) example, if a toggle switch is used, the words "Flip up" or "Flip down" are preferred over less specific words such as "Manipulate." In addition, the analyst should record the position to which a switch is to be set or some other indication of response adequacy.

The fourth column, "Cue for initiation or completion of activity," is used to indicate the cue that tells the operator when to begin a step or the cue that tells him when he has successfully completed a step. In most cases, this information is found in the "Activity" column, but it is used in this format to remind the analyst to pay particular attention to these cues. Errors can result if the design of equipment or procedures does not provide good cues. Misleading or incomplete cues can result in discontinuities in carrying out a task, with the result that some step in a procedure may be omitted.

The last column, "Remarks," is used for relevant information not covered in the other four columns. The completed task description provides the information for the analytic part in which the identification of ELSs and APSs is made.

The task description will be fairly gross at first. Identification of the tasks is probably all that should be done initially; i.e., a task listing. This will enable the analyst to relate that which people have to do to the various functions defined in Step 1 in the MMSA. It may be useful to key the tasks to flow charts developed from Step 1. More detail in the task description will be possible as a system design becomes more definitive. When all of the procedural steps have been recorded, they can serve as entries to procedures manuals and training courses.

One useful aid to task description in NPP studies is link analysis. This tool is often used in laying out work places and job operations, and can be used to study the interactions among people in an existing plant. A link analysis depicts the pathways among different parts of a system as they are generated by people walking about and communicating.

Figures 3-2 and 3-3 (pp 3-28 and 3-29) show a link analysis for the reactor engineer for an afternoon shift at a Swedish BWR (Axelsson and Lundberg, 1975). Additional link analyses from the same study showed that the reactor engineer, turbine engineer, and shift supervisor had to spend much time looking at displays not visually accessible from their normal working positions. This kind of analysis can suggest improvements for future designs, and, for human reliability analysis purposes, can provide an understanding of the difficulties and inconveniences that influence human performance. Procedures for performing link analyses are described in Chapanis (1959, pp 52-61) and McCormick (1975, pp 293-298).

Another useful technique for outlining operating time and personnel interactions is called the operational sequence diagram (Brooks, 1960 and Kurke, 1961). The operational sequence diagram displays information-decision-action sequences in a man-machine system. It can be used in preparing time-sequence process charts or spatial flow charts. This technique involves some symbolic shorthand, but the number of symbols to be learned is not excessive. The main advantage of these diagrams is that they outline essential interactions among operators, work stations, items of equipment, and time.

Step 5 - Analyze the Jobs and Tasks to Identify Error-Likely Situations (ELSs) and Other Problems

In the analytic part of the task analysis, each human action is analyzed to identify ELSs arising from equipment design features, methods of use, methods of training, and the skill levels of the people in the system. There are no hard-and-fast rules for making these determinations. The validity of the task analysis will depend upon the skill of the analyst in assuming the role of the operator so that he can understand the actual and potential problems in each task.

Even the best analyst cannot identify all possible modes of human response. In terms of our human error classification, he will not be able to predict all errors of commission and all possible extraneous acts by plant personnel, such as the use of a candle to check leaks in the negative pressure containment building (the Brown's Ferry Incident). Still, given sufficient time, a skilled analyst can identify most of the important tasks to be performed in a system, and most of the ways in which errors may be committed.

The "Analytical" half of the format in Figure 4-1 indicates the kinds of factors to consider in identifying an ELS. The listed factors are under four broad headings: (1) Scanning, perceptual, and anticipatory requirements; (2) Recall requirements (long-term or short-term memory) and initiating cue (present, absent, or poor); (3) Interpreting requirements; and (4) Manipulative problems. The terms are self-explanatory, are relevant to different situations, and may be used without modification for NPP analyses. The column headings should be regarded as suggestions only -- any unique problem should be listed regardless of the column headings presented here.

In use, the analyst makes entries in this half of the form only when he identifies an ELS. For example, assume that he finds an ELS in Step 3. At that point in the form he notes the basis for his judgment of an error-likely task. Referring to the factors in the columns, an ELS exists when the discriminating, recalling, interpreting, inferring, decision-making, or manipulating processes demanded of the operator are likely to exceed his capacity. The potential errors can be errors of omission or commission, extraneous acts, or sequential or time errors.

The analysis is done for each task or step in a task to determine those PSFs that seem likely to result in errors. It must be determined whether there are any conflicts between the external PSFs and the internal PSFs, since such conflicts can be expected to result in errors. Chapter 3 lists examples of external PSFs which are not compatible with various human attributes and therefore result in lowered human reliablity. Some conflicts between the external and internal PSFs can cause psychological or physiological stresses. If the level of stress is high, the performance of a person in the system will probably deteriorate. On the other hand, if the level of stress is too low (as with monotonous work), alertness may be degraded and signals may not be noticed soon enough.

In summary, we define error-likeliness in terms of those PSFs in a task that are incompatible with the capabilities and limitations of the intended performer of the task. The task analysis will indicate if human reliability can be improved by changing any PSF.

Whether or not it is important enough to the system to warrant changing the design is another matter. The object of task analysis is to identify potential sources of error regardless of their impact on the system. Other steps in the MMSA take the consequences of error into account, as discussed later.

Several publications are available to assist the analyst in identifying ELSs. The most concise document is MIL-STD-1472B, Military Standard, Human Engineering Design Criteria for Military Systems, Equipment and Facilities, U.S. Dept. of Defense, Wash., DC, 15 May 1970, with Notice 1 dated 10 May 1976 and Notice 2 dated 10 May 1978. This set of standards was developed by practicing ergonomists in the U.S. Army and in U.S. industry, and adopted by all of the U.S. military services. These standards are not absolute, but their acceptance in the design of NPPs would materially improve human reliability as it has in countless complex military systems. They represent sound human engineering practices to be followed unless specifically contraindicated by other aspects of the system. A companion document, MIL-H-46855B, Military Specification, Human Engineering Requirements for Military Systems, Equipment and Facilities, U.S. Dept. of Defense, Wash., DC, 31 January 1979, defines the general requirements for incorporating ergonomics considerations in the league of systems. Other documents useful in identifying ELSs are the revised edition of the Human Engineering Guide to Equipment Design (Van Cott and Kinkade, 1972), the revised edition of the Human Engineering Guide for Equipment Designers (Woodson and Conover, 1964), and two textbooks: Human Factors in Engineering and Design (McCormick, 1975) and Ergonomics: Man in His Working Environment (Murrell, 1969). These documents provide much of the rationale and data behind the standards in MIL-STD-1472B. The two volumes by Edwards and Lees (1973 and 1974) constitute the best available description of operator roles in complex industrial processes analogous to those in NPPs. Finally, the 3-volume study of human factors problems at the TMI-2 plant by Malone et al (1980) not only includes examples of poor ergonomics but describes the kinds of studies necessary to identify such problems.

Although the above documentation will be useful, the best way for the analyst to determine which human processes and actions will be employed in performing each task is to do the tasks himself, using whatever written procedures are available. Then he should observe and interview operators who perform the tasks. Since highly skilled operators can make even poor designs look good, it is necessary to include talk-throughs or walk-throughs in the observation of the operators at their jobs. This technique involves the introduction of pauses in the human actions while the operator explains what he is doing and his mental processes. The analyst should observe the operations being performed at their normal speed until he develops a sense of familiarity. Then he should ask the operator to slow down his work activities and explain what he is doing, and why. As he performs the tasks himself and interacts with the operators, the analyst will develop familiarity with the system hardware and procedures. This is the period when he will obtain the data for the analytical half of the task analysis format. Table 4-2 is a general checklist that can be used during this period.

Another useful technique is to have the operator talk-through hypothetical, but realistic, emergency problems. In the WASH-1400 study, this technique was employed to find out how much operators knew about responding to certain emergency conditions, and what supporting provisions were made for these responses in the design of equipment and written procedures. Talk-throughs can also reveal the mental model the operator has of the plant and its processes.

Experienced people on a job can rightly be regarded as subject-matter experts. They know more about the intricacies and difficulties of their tasks than anyone else. The importance of interviewing such experienced operators is that they are invaluable in identifying problem areas in tasks. They can describe errors they have made or have seen others make (including no-cost errors), and can offer opinions on the underlying PSFs related to these

Table 4-2. A Checklist for Evaluating Task Error-Likeliness During Observation or Performance of Tasks

1. The cue or sign that tells the operator to begin each task and each activity in a task is simple and unambiguous:
    - a. No inconsistencies, gaps, or misleading information that could result in errors of judgment
    - b. No competing activities that could interfere with perceiving the cue or with responding in a timely manner
2. The cue or sign points to the correct task or activity <u>only</u>.
3. The task or activity is easy to do:
    - a. No special problems in the scanning, anticipatory, or other perceptual requirements; in long-term or short-term memory requirements; in interpreting and decision-making requirements; or in manipulative requirements
    - b. No special problems with competing activities or past experience
4. The completion cue or sign for the task or activity is simple and unambiguous:
    - a. No misleading feedback to the operator
    - b. No special problems with competing activities
5. The completion cue or sign for one activity in a task cannot be misinterpreted as signaling the completion of the entire task

errors. These subject-matter experts can also describe close calls and, in general, what errors are possible and likely.

In performing a task analysis, the underlying behavioral processes required in each task must be identified. As shown in Figure 3-1 (p 3-2) it is convenient to think of these processes as consisting of input variables, mediating variables, and output variables, a conventional trichotomy in psychology. Referring to Figure 3-1, signs (usually visual or auditory displays) feed information to the senses of the operator. These are the inputs that define the discriminations he must make. The discriminations are determined both by the features of the task to be sensed and by the individual characteristics of the operator -- his sense organs, past training and experience, any ongoing activities that compete for his attention, his emotional state, and so on. In our post-WASH-1400 studies, we have found that much reliance is placed on these internal PSFs to make up for inadequate human factors engineering of the job situation. In a well-designed system, the equipment, procedures, etc, do not place undue demands or reliance on operator characteristics.

The motor responses are those outputs of the operator with which he performs some element, or step, in a task. (His responses may or may not be appropriate.) When using the event trees described in the next chapter, each task element is treated as either a success or a failure in terms of system requirements. When this either/or distinction is not appropriate, different degrees of success or failure can be treated as different events. For example, assume that there is some probability that an operator will leave a manual valve in some position between open and closed. Although there are an infinite number of in-between positions, for any practical application, this number can be reduced to one (or a very few) in terms of system effects, and each of these positions would be treated as an event. Such reduction is necessary to keep the analysis manageable.

Task analysis is applicable to continuous as well as discrete tasks (Miller, 1953b; and Meister and Rabideau, 1965). Examples of the former include the analysis of the in-flight functions of an aircraft pilot, the tracking employed in air-to-air flexible gunnery, and the tracking tasks in operating a continuous strip mill. In such cases, the continuous nature of these tasks is described as a series of discrete task elements. This type of abstraction is true of human performance modeling in general, and is used here. Although continuous tasks can be used directly as entries in task analysis or performance modeling, the solution methods are cumbersome, and are much simplified if the variables are treated as discrete values -- the differences in accuracy are negligible for practical work.

The mediating processes are the internal responses of the operator, such as thinking, deciding, and worrying. These processes constitute the bridge between the inputs to the operator and his outputs. Although not directly observable, the processes can be inferred by attention to the physical features of the task and the known needs, capabilities, limitations, and motivations of the operator. To understand and identify these processes, interviews with the operators are the best source of data. If you want to know what is going on in the mind of an operator, ask him. Obviously, this technique is subjective; the operator may not really know why he does certain things, or he may wish to deceive the analyst deliberately. Finally, thinking about what he does may actually change what he does. Still, invaluable information can be obtained by interviewing operators as they work. (For an example of interviewing to study the mental processes involved in electronics troubleshooting, see Rasmussen and Jensen, 1974.)

Figure 3-1 (p 3-2) shows that the operator's output produces system results and that information about these results is fed back to the operator's sense organs via displays. Thus, the man-machine system is a negative feedback system in that information about the output can be compared to some standard (i.e., a desired output), and the operator can take corrective action to minimize the system error.

The most difficult part of a task analysis is to identify the possible unplanned modes of operator response. It is not too difficult to set down in sequence the tasks and steps in these tasks for well-defined jobs with more or less invariant sequences of actions. Thus, a task description of a calibration technician's job is straightforward. One can readily set down what leads to what, and the identification of potential errors is not difficult.

In other jobs in the NPP, tasks may not be as well-defined, and variable sequences may be common. Even in "routine" tasks, one must consider how the operator might deviate from the anticipated routine. In less-structured tasks, such as those involved in responding to unusual circumstances, the job of the analyst is more difficult. He must identify where and how the operator's responses to unusual events might create more demands on the system instead of correcting the situation. This aspect obviously requires that the analyst have a high level of knowledge of the job he is describing and analyzing.

Step 6 - Estimate the Likelihood of Each Potential Error

Steps 6, 7, and 8 of the MMSA provide an estimate of the importance of each ELS identified in the task analysis. The importance of an error is a function of its frequency, probability of recovery, potential consequences, and the costs of fixing the underlying ELS.

In the most quantitative form of human reliability analysis the frequency of human error is converted to a probability estimate. Chapter 20 summarizes the human error probabilities that are relevant to NPP operations, and Chapter 5 shows how these HEPs are handled in the analysis. These HEPs generally have widely spaced uncertainty bounds to allow for individual differences as well as other sources of uncertainty. If more precise data are available, they should be used instead. For some purposes, the upper bound of the uncertainty range may be used as a "worst" estimate. If an entire analysis is based on these high estimates of HEPs, the result will be a worst-case analysis, an example of which can be found in Chapter 21.

The context of any event must be considered in order to estimate the probabilities of human events. For human events, interaction (dependence) is the rule rather than the exception. No procedure exists for mechanically combining basic error probability data into total estimates of task failure probabilities. The analyst must use the information from the task analysis to determine the important PSFs in deriving the probability estimates for each task or subtask. The examples of probability tree diagrams throughout this handbook illustrate the types of judgments used to derive these estimates.

The purpose of breaking down a task into inputs, mediating processes, and outputs is to obtain smaller bits or elements of behavior that can more readily be combined with available data. This task decomposition makes it easier to identify all of the PSFs that influence the reliability of a human task and to evaluate the adequacy of available data for assessing task reliability.

To cite a simple example, suppose there are data available on the reliability with which experimental subjects read 6-digit numbers, but the task the analyst is studying involves reading 3-digit numbers. Other things being equal, the error probabilities based on reading 6-digit numbers would overestimate the error probability for reading 3-digit numbers. Therefore, the analyst would not want to use the available data without some adjustment -- which might be obtainable from other experiments.

The above example is simplistic. In a real analysis, the other PSFs in the experiment would have to be compared carefully with the PSFs in the task being analyzed. We have never found reports of experimental conditions that were identical to the task conditions of interest.

Extrapolations are the rule rather than the exception. Occasionally, if failure probability data exist for tasks similar to the tasks of interest, the decomposition of tasks into smaller bits of behavior may be unnecessary. However, the decomposition is often worthwhile to ensure that all major possibilities for error have been assessed.

When estimates of HEPs for individual behavioral units have been obtained, they can be combined into estimates of HEPs for larger units of behavior corresponding to entire tasks or groups of tasks. In this recombination, the estimated HEPs for small behavioral units often must be modified in consideration of their interdependences. Usually, the estimate of an HEP for a large unit of behavior is not merely the addition of a group of estimated HEPs for smaller units of behavior. Before this combination can take place, the interdependences must be considered and the error contributions modified as described in Chapter 7.

Step 7 - Estimate the Likelihood that Each Error will be Undetected (or Uncorrected)

Other things being equal, the smaller the likelihood that some error will be detected before it causes undesirable consequences, the more important the error. Some errors will be detected by the person who makes them. (These are examples of recovered errors.) Other errors may be detected by inspectors or by subsequent testing and use of system components, but some errors may not be detected until unwanted consequences to the system have occurred. In a human reliability analysis, recovery factors (Chapter 15) are used to modify the HEPs, since the interest is in estimating the joint probability that an error will be made and will not be recovered (i.e., the probability of an unrecovered error).

Step 8 - Estimate the Consequences of Each Undetected (or Uncorrected) Error

The consequences of an error obviously define another aspect of the error's importance. In a well-designed system a single uncorrected human error rarely causes serious degradation. Although there have been such cases, normally there is sufficient redundancy in the system such that these errors will not result in serious consequences. For example, although a calibration technician may miscalibrate a single setpoint for some temperature sensor, there will be other sensors that will indicate a disagreement. Appendix II of WASH-1400 shows how various unrecovered error probability estimates were incorporated into the ESF system fault trees.

The usual procedure in human reliability analyses is to perform a separate analysis for each system consequence of interest. Generally, quantification of the relative importance of each consequence is not part of the human reliability analysis. For example, separate human reliability analyses would normally be done for the influence of human errors on the risk of personnel injury and on the risk of some economic loss. That is, one would not try to place these two system consequences on a single continuum of loss.

In reliability assessments it is often of interest to learn how the probability of failure of a system involving many human tasks would change if different estimated HEPs were used for the individual tasks. Such an assessment is called a sensitivity analysis. It is very useful in human reliability analysis because estimates of HEPs are ordinarily made with uncertainties larger than those assigned to estimates of failure probabilities of other system components. Sensitivity analysis was used in some of the early analyses in WASH-1400. It was discovered that the probabilities of failure of some large subsystems in NPPs were insensitive to substantial variations in estimated HEPs as well as to variations in assumed distributions of HEPs. (Chapters 7 and 21 present some examples of sensitivity analysis.)

Step 9 - Suggest Changes to the System

This step is primarily related to the use of MMSA as a design tool. Most consideration is given to those potential errors with a high combined probability of (1) occurring, (2) going undetected or uncorrected, and (3) causing an unacceptable system consequence. Thus, a task with a high error probability may be less important to system success than some other task with a lower error probability. For example, if the first task has good recovery factors and the other one does not, the latter task may have more potential for degrading the system.

Decisions to incorporate ergonomics changes in a system often require trade-offs of various criteria and costs. Although the issue of such trade-offs is outside the purview of this handbook, Table 4-3 lists some important system criteria. (These criteria are not listed in order of importance.) It is clear that costly changes should not be recommended solely because a design may deviate slightly from the optimum. However, gross deviations from optimum

Table 4-3. System Criteria for Trade-Off Considerations*

1. Speed (mean and variability)
2. Accuracy (constant error) and precision (variable error)
3. Dependability (maintainability and reliability -- including confidence in self-check)
4. Adaptability (of equipment to changes in requirements, equipment design, or operating conditions)
5. Mobility (including dispersal requirements)
6. Graceful degradation (ability to continue to operate although at substandard levels of performance)
7. Equipment test life (need to avoid adverse effects of confidence or other testing)
8. Completeness or exhaustiveness (the proportion of system parameters that must be measured)
9. Personal involvement (extent to which personnel identify themselves with their tasks or are aware of system operation)
10. Personnel hazard and risk of equipment damage
11. Delivery schedule (time for system to become operational)
12. Equipment weight and/or volume
13. Training costs (personnel, time, facilities)
14. Manning level and quality
15. Development costs
16. Logistics costs and policy (pipeline and spares provisioning policies)
17. Equipment unit cost in production (including spares)
18. System environment (ability to operate under various climatic, terrain, socio/psychological, political and other conditions)
19. Selling costs (including advertising)
20. Aesthetic attractiveness
21. Effects on environment (ecological considerations)
22. Costs of employee dissatisfaction (indifferent and slow work, absenteeism, turnover, grievances, strikes, sabotage)

*Modified from Swain and Wohl, 196

ergonomics design principles have resulted in errors in large numbers and varieties of man-machine systems. Kemeny (1979) and Sugarman (1979) describe several such deviations that contributed to human errors in the TMI accident. Based on experience with other complex man-machine systems, the incorporation of standard ergonomics principles in the early design stages of TMI-2 would not have been costly.

A useful rule is to follow optimal ergonomics principles unless there are overwhelming reasons to disregard them. The consideration of such principles early in the design phases of a system will usually allow for the incorporation of optimum human factors design at minimal cost.

Using the human reliability analysis as a guide, suitable design changes can be developed to reduce the probability of system degradation. A candidate design change may reduce the probability of an error (e.g., by reducing the number of opportunities to make the error); it may increase the likelihood that an error will be detected or corrected (e.g., by providing better feedback to the operator or by adding an inspection step); or it might involve some provision for the system to tolerate an error.

The design changes may address any of the PSFs associated with the potential error. Sometimes the design change is as simple as changing the scale on a meter, the color of an indicator lamp, or the size of a handle. At other times the design change might be more costly, such as a rearrangement of controls, displays, or panels. At times a desired design change may not be feasible for some reason, and the only alternative may be to provide human redundancy by assigning an extra operator when certain procedures are to be carried out; e.g., assigning a second control room operator who is dedicated to maintaining adequate levels of coolant in the primary and secondary loops, a practice initiated by some plants after the TMI accident.

Obviously, the decisions are not always as simple and clearcut as the above examples suggest. The final decision will have to be made on the basis of acceptable risk and cost-effectiveness.

Step 10 - Evaluate the Suggested Changes (Repeat Steps 1 through 9)

Finally, each suggested change to the system must be reevaluated by repeating most of the above steps. Thus, MMSA is iterative -- the steps are repeated until the estimated human error contribution to system degradation has been reduced to some tolerable level in terms of system effects and costs of the changes. The contribution of human error may be reduced either directly by improvements made to reduce error frequency, or indirectly by design changes that will tolerate human errors. The best solution is obviously situation-specific.

CHAPTER 5. THE HUMAN RELIABILITY MODEL

The human reliability model described in this chapter is an extension of human reliability studies made at Sandia National Laboratories in the early 1950s (Swain, 1964b). It was first used to estimate the quantitative influence of first-order human failure terms on the reliability of nuclear weapon systems and components. In the early 1960s the model was expanded and refined to permit more detailed consideration of the human component in system reliability. Subsequent development of the model has included its applications to a large variety of classified systems, to the U.S. NRC's Reactor Safety Study (WASH-1400), and to subsequent human reliability problems in NRC-supported work. Some of these latter applications are presented in Chapter 21.*

Most of the applications of the human reliability model and method described in this chapter have involved estimates of the probabilities that system-required tasks will be executed correctly within specified time limits. Applications of the model to the prediction of the effects of extraneous acts have been limited to worst-case analyses, as described in Chapter 21.

There are other human reliability methods and models, but none of them has had any extensive practical application, and some of them depend on nonexistent data or do not result in estimates of HEPs. Two extensive reviews have been made to Meister (1971) and Embrey (1976). The former volume has a critical review of 22 human reliability analysis models.

---

*Those interested in the history and use of this model can refer to: Rook, 1962, 1963, and 1964; Swain, 1963a and b, 1964a and b, 1967a and b, 1969a and b, 1971, 1974a and b, 1976, and 1977a and b, 1980b (Ch. VIII); Rigby, 1967; Rigby and Edelman, 1968a; Rigby and Swain, 1968; Swain and Guttmann, 1975; and "Human Reliability Analysis", Section 6.1, Appendix III - Failure Data, 1975, WASH-1400.

Description of the Model

At the Sixth Annual Meeting of the Human Factors Society in November, 1962, Swain introduced the acronym THERP to designate the human reliability model and method that he and Rook had developed at Sandia National Laboratories (Swain, 1963a). THERP stands for Technique for Human Error Rate Prediction. Until publication of this handbook we used the expression human error rate (HER) interchangeably with human error probability (HEP). For reasons stated in Chapter 2, we have dropped the term HER in favor of HEP. However, since the acronym THERP is now well known, we retain it despite its use of the term human error rate.

The following is a revised definition of our human reliability technique:

> THERP (Technique for Human Error Rate Prediction) is a method to predict human error rates [i.e., human error probabilities] and to evaluate the degradation of a man-machine system likely to be caused by human errors alone or in connection with equipment functioning, operational procedures and practices, or other system and human characteristics that influence system behavior.

The model uses conventional reliability technology with modifications appropriate to the greater variability and unpredictability of human performance, as compared with that of equipment performance. The steps in THERP are similar to those in conventional reliability analysis, except that human activities are substituted for equipment outputs. The steps are to:

1. Define system failure(s) of interest. These pertain to system functions which may be influenced by human errors and for which error probabilities are to be estimated.
2. List and analyze the related human operations. This step is the task analysis described in Chapter 4.
3. Estimate the relevant error probabilities.

4. Estimate the effects of human errors on the system failure events. This step usually involves integration of the human reliability analysis with a system reliability analysis.
5. Recommend changes to the system and recalculate the system failure probabilities. (The procedure is iterative.)

The above five steps typify the use of human error analysis as a tool in system design. For assessments only, Step 5 is not required.

In using THERP, the primary interest is in estimating the following parameters, especially the first three:

1. Task Reliability -- Task reliability is defined as 1.0 minus the estimated probability of task failure. For each task we determine the probability that it will be completed successfully (within some period of time, if time is a requirement). The tasks are identified in the task analysis and an estimate is made of the failure probability for each task. Effects of extraneous actions must also be considered.
2. Error Correction -- This is the probability of detecting and correcting incorrect task performance in time to avoid any undesirable consequences. In any man-machine system there are usually several recovery factors; e.g., checks by other people (inspectors) which increase the probability of detecting errors before they affect the system.
3. Task Effects -- This is the probability that incorrect and uncorrected task performance will result in undesirable consequences to a system. A separate calculation is made for each system consequence of interest. Therefore, one may estimate the effects of the same human errors on more than one system outcome.

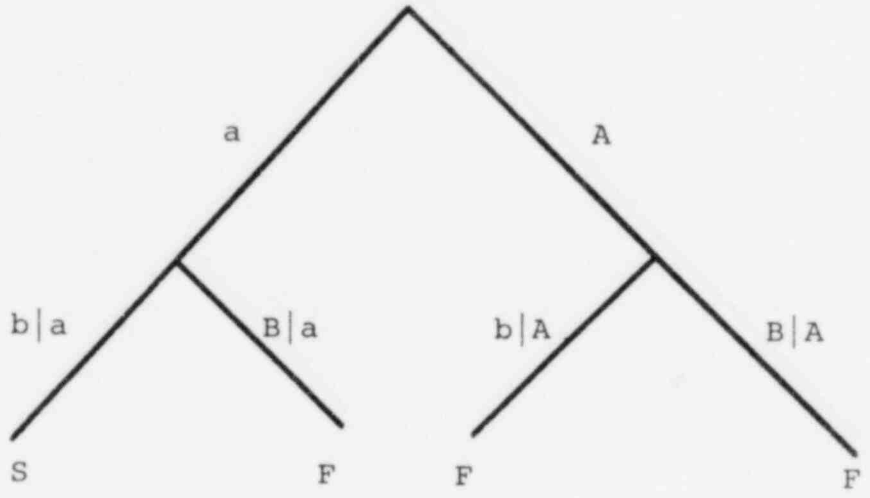
4. Importance of Effects -- The importance of the undesirable effects to a system in terms of cost or other criteria should be considered. Generally, no attempt is made to quantify this parameter; it is often a value judgement made by persons in authority.

THERP is used to generate quantitative estimates of the first three parameters based on the dependences among human performance, equipment performance, other system events, and outside influences. Thus, estimates of HEPs for all but an initiating task represent conditional probabilities.

Probability Tree Diagramming

The basic tool of THERP is a form of event tree called the probability tree diagram (Figure 5-1), in use at Sandia National Laboratories since the 1950s (Müller, 1964). Limbs in the probability tree diagram show different events as well as different conditions or influences upon these events. Therefore, the values assigned to all events depicted by the tree limbs (except those in the first branching) are conditional probabilities. The first limbs may also be conditional probabilities if they represent a carryover from some other tree.

Table 5-1 presents the symbology used with the limbs of the event tree. Note that a letter can have more than one meaning depending on whether it is in quotes, capitalized, or lower case. A capital letter in quotes represents an event or task. For example, Task "A" in Figure 5-1 might represent the task of a calibration technician setting up his test equipment before calibrating some sensor. The lower case letter a represents the statement that the technician has correctly set up his test equipment, and also stands for the probability of correct performance. The capital letter A represents the statement that the technician has incorrectly set up the test equipment, and also for the probability of incorrect performance.



a = probability of successful performance of Task "A"

A = probability of unsuccessful performance of Task "A"

b|a = probability of successful performance of Task "B" given a

B|a = probability of unsuccessful performance of Task "B" given a

b|A = probability of successful performance of Task "B" given A

B|A = probability of unsuccessful performance of Task "B" given A

$$Pr[S] = a(b|a)$$

$$Pr[F] = 1 - a(b|a) = a(B|a) + A(b|A) + A(B|A)$$

Figure 5-1. Probability Tree Diagram of Hypothetical Calibration Job

5-1. Symbology Used in Human Reliability Model

Symbol	Meaning
Capital English letters in quotes; e.g., "A"	1. The human action itself; e.g., Task "A."
Capital English letters; e.g., A (Except F and S)	1. Incorrect performance of a human action. 2. Probability of incorrect performance of a human action.
Capital letter F at the end of a path through an event tree	1. Failure, for the application in question.
Capital letter S at the end of a path through an event tree	1. Success, for the application in question.
Lower case English letters; e.g., a (except i and r)	1. Successful performance of a human action. 2. Probability of successful performance of a human action.
Lower case English letters, i and r	1. i^{th} or r^{th} task.
Lower case underlined English letter, $\underline{n}$	1. The number of events or tasks, not to be confused with n, which indicates the successful (or probability of successful) performance of Task "N."
Capital Greek letters; e.g., Δ	1. The estimated probability of nonoccurrence of some system event (not a human event).
Lower case Greek letters; e.g., δ	1. The estimated probability of occurrence of some system event (not a human event).

The letter b represents the statement of correctly performing Task "B," the second task performed, and it also stands for the probability of correct performance. Task "B" might be the calibration of the sensor mentioned above. The letter B stands for incorrect performance as well as for the probability of incorrect performance. The dependences between Tasks "A" and "B" are represented by the symbols $b|a$, $B|a$, $b|A$, and $B|A$. Normally, the conditional relationships are understood, and the letters b and B are usually written without the conditional qualifications.

In Figure 5-1, only the *complete-success path*; i.e., $a(b|a)$, is designated as ending with S, which stands for some success path of interest. All the other paths end in F, which stands for failure. Thus, this tree as drawn indicates that the only success path is one in which *both* tasks are correctly done; i.e., the calibration technician correctly sets up his test equipment and also correctly calibrates the sensor. For other problems, the interest might be in performing either task correctly, and any path other than $A(B|A)$, the *complete-failure path*, would be considered a success path. It is the application that determines which paths through the tree are considered success paths or failure paths.

In Figure 5-1 the limbs in the tree represent a binary decision process; i.e., correct or incorrect performance are the only choices. Thus, $a + A$ must equal 1.0, and $b + B$ must equal 1.0. At every binary branch the probabilities of the two branches sum to 1.0. In other probability tree diagrams there may be more than two limbs at a branching to represent different conditions or events, or different levels of correctness or incorrectness, but in all cases the probabilities of the limbs at any one branching must sum to 1.0. As with any probability tree diagram, the sum of the probabilities at the terminals of all paths also must sum to 1.0. Thus, in Figure 5-1, $(a \times b|a) + (a \times B|a) + (A \times b|A) + (A \times B|A) = 1.0$.

Usually the limbs in an event tree will represent correctness or incorrectness of system-required tasks, but they can also denote extraneous acts that can be anticipated and that can adversely affect the system. In the human reliability analyses done for WASH-1400, for example, besides depicting the selection of certain correct switches, limbs were also used to depict the selection of nearby incorrect switches that could cause serious problems in coping with a LOCA if inadvertently selected. Thus, limbs in event trees can be used to represent plausible and important extraneous actions. Of course, not all extraneous actions can be identified in advance. Although most such actions are unimportant in terms of system consequences, it is always possible that some important extraneous action will be overlooked in a human reliability analysis. The more detailed the task analysis behind the event tree, the greater the likelihood of identifying the important, plausible extraneous actions.

After the probability tree diagram is drawn, the mathematics are simple. When the estimates of the conditional probabilities of success or failure of each limb in the tree have been determined, the probability of each path through the tree is calculated by multiplying the probabilities of all limbs in the path. This does not correspond to the simple multiplicative model; that is, the multiplication of task probabilities without the assumption of any task dependences (the multiplication of unconditional probabilities). The use of conditional probabilities takes into account the interdependences among the limbs in the tree, and no errors will result from the use of this simple mathematical approach. Errors would arise from incorrectly estimating the BHEPs or CHEPs for the tasks represented by a limb.

The probability tree diagram starts with any convenient point in a system procedure and works forward in time. This procedure enables the user to analyze what leads to what, and allows him to identify the important events affecting human performance. In the fault tree approach one starts with some fault and works backwards in time. Thus, the event tree is an inductive model, the fault tree a deductive model.* As with fault trees, boxes with short descriptions of events instead of symbols may be used in an event tree to assist the reader. Also, as with fault trees, one limb in a probability tree diagram may stand for another entire tree. Nielsen (1971 and 1974) has developed a combination of event trees and fault trees called Cause/Consequence Diagrams; this combination has some of the advantages of both schematic techniques.

In WASH-1400, the system reliability analysts used a different format for event trees (see Figure 5-2). In this type of event tree, the progression is from left to right, and the failure limbs are the lower limb in each branching. Whichever format is used for an event tree is a matter of convenience for the analyst. We prefer our top-to-bottom progression, which is the one we (the human reliability analysts) used in WASH-1400.

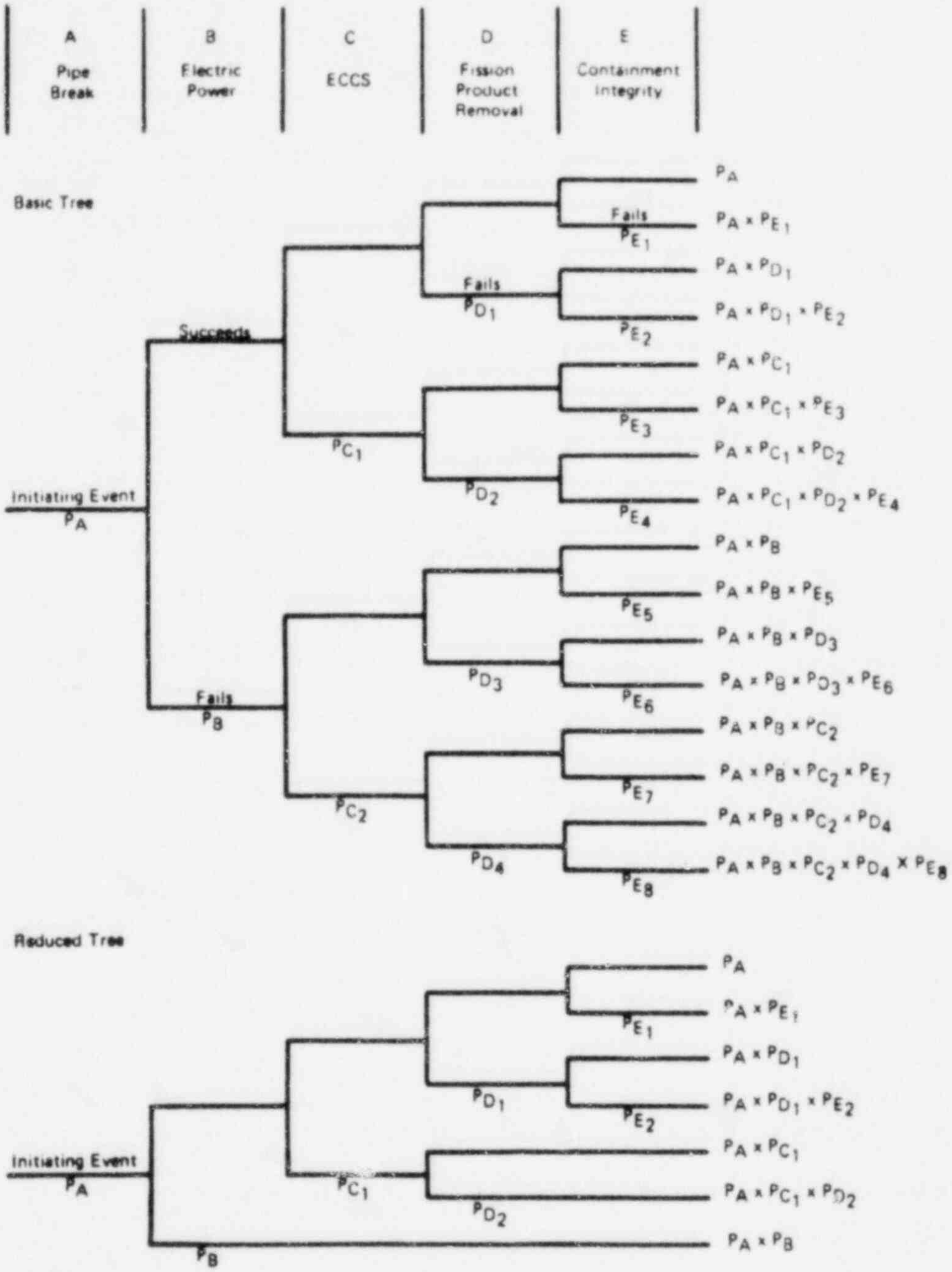
Conditional Probabilities

Use of the probability tree diagram is based on the assumption that the estimates associated with the limbs are conditional probabilities. Considerable error in estimating the probabilities of paths through the tree can occur unless conditional probabilities are used. If independence of acts is naively assumed, the failure contributions of people to systems can be significantly underestimated.

As an example, let us take the tree in Figure 5-1, but now assume that the only failure path is A(B|A); i.e., failure occurs only if Tasks "A" and "B"

---

*See WASH-1400, Appendix II - Fault Trees, pp II-27 to II-35 for a description of the use of fault trees in analyzing NPP operations and for the incorporation of estimated HEPs into the trees.



Note Since the probability of failure, P is generally less than 0.1, the probability of success (1 P) is always close [illegible] the probability associated with the upper (success) branches in the tree is assumed to be 1.

Figure 5-2. Simplified Event Trees for a Large LOCA (from Fig. 4-4 of the Main Report to WASH-1400) are both performed incorrectly. All the other paths lead to system success. Assume that Tasks "A" and "B" are the opening of two valves after maintenance, and that the system will function properly if either valve is opened. Further assume that the valves are next to each other, so that if the operator forgets to open one valve, he is 100% sure to forget to open the other valve; i.e., there is complete dependence between the two acts. Also assume that the BHEP of forgetting to open a valve of this type is 10^{-2}. If we naively assume independence between the opening operations for these two valves, we will get an unrealistically low failure probability estimate of $A \times B = 10^{-2} \times 10^{-2} = 10^{-4}$. However, if we correctly assign the probability of 1.0 to $B|A$, our estimate of the failure probability is $A(B|A) = 10^{-2} \times 1.0 = 10^{-2}$. This example illustrates the importance of dependences among the events in a reliability tree diagram. Chapter 7 presents the dependence model in detail.

Distribution of Human Performance

In human reliability analysis, probability distributions of human performance are usually represented by one or a few values. The distribution of human performance on a task is usually ignored, and some estimate of the central tendency of a population is used; i.e., single point HEPs are used. This approximation is adequate for most applications.

For cases in which human performance would be greatly influenced by different levels of some PSF, two or three levels of that PSF are used (rather than the entire distribution), as in the calibration example in Chapter 7 (p 7-19). In such cases, the tasks to be performed under different conditions are represented by different branches in the probability tree diagram, and the appropriate conditional estimates are assigned to each limb in the tree (Swain 1963b, 1974a, and 1976).

Sometimes it is necessary to estimate the distributions of human performance because the human reliability analysis is to be a part of some system reliability analysis in which computer procedures will sample from specified distributions rather than use specified points. For example, in WASH-1400 a Monte Carlo procedure was used to sample from a distribution of probabilities of both equipment and human failures, and a range ratio was take [illegible] represent the ratio of the 95th percentile HEP to the 5th percentile HEP. In general, the 95th percentile HEP was estimated as either 3 times or 10 times the nominal HEP (i.e., the estimate of the central tendency of error probabilities for a task), and the 5th percentile HEP was estimated as the nominal HEP divided by either 3 or 10. This treatment is basically the same as that used in this handbook, which considers the upper and lower uncertainty bounds to represent, respectively, 95th and 5th percentile HEPs.

The same general assumptions about the shape of the distributions of human events made in WASH-1400 are used in this handbook -- we assume log-normal distributions or roughly equivalent distributions. As discussed in Chapter 16, "Distribution of Human Performance and Uncertainty Bounds," we have reevaluated the assumption of a log-normal distribution made for the WASH-1400 analysis and find no reason to change it.

Use of Model Outputs in System Reliability Studies

The outputs from the human reliability model will consist of estimates of the success or failure probabilities of human actions or tasks, modified by system contexts and events. In a reliability analysis of an NPP system, the estimates of human reliability can be included in fault tree analyses, provided that the dependences among human events are also incorporated. Thus, the outputs of the human reliability model are estimates of human component reliability and can be handled as are estimates of other system components.

In WASH-1400, for example, fault trees showed the relationship of some particular component to the availability of some engineered safety feature. Figure II 5-45 (p II-303 in Appendix II, Fault Trees, in WASH-1400) shows that an operator might forget to open MOV-1866E after a monthly flush of the High Pressure Injection System (HPIS) discharge lines. The estimated probability of an unrecovered human error is 3×10^{-4}. If that human error occurs, or if any one of three mechanical failures occurs, the HPIS will fail to deliver sufficient water through the HPIS discharge line to Cold Leg 2. Thus, in the reliability analysis human errors are treated in the same manner as are failures of mechanical components.

To arrive at the 3×10^{-4} error probability estimate in the fault tree, it might be necessary to construct a probability tree diagram to ensure consideration of all the events related to the failure to open MOV-1866E. However, as often occurs, this failure event was so simple that a tree diagram was not necessary. That is, for the error to occur, an operator has to fail to initiate the task of reopening the MOV or fail to operate the MOV switch correctly in the control room, and someone else has to fail to inspect the switch or fail to note that it is closed when he makes the inspection.

In the WASH-1400 analysis, the estimated 3×10^{-4} unrecovered error probability was derived by using the 3×10^{-3} estimated probability of an error of omission for items that are embedded in a procedure (Table III 6-1 in WASH-1400). This estimate was modified by assuming a 10^{-1} probability that either the inspection would not be carried out or the inspector would fail to notice that the switch was in the incorrect position. Thus, $\Pr[F] = 3 \times 10^{-3} \times 10^{-1} = 3 \times 10^{-4}$. The 10^{-1} estimate included the dependence between the operator and the inspector. For the WASH-1400 analysis, this gross estimate was adequate.

More precise estimates can be obtained by performing a detailed task analysis of each specific task, constructing one or more probability tree diagrams to indicate the success paths and the plausible failure paths, and assigning the appropriate estimates of conditional probabilities of success and failure to the branches.

The most obvious sources of information for the probability tree diagram for the tasks to be performed are the written procedures for the tasks themselves and the plant's operating rules for verifying of the correctness of the tasks performed. However, in many instances, basing the event tree solely on written procedures and plant operating rules is not likely to result in a complete and accurate picture of the typical human actions. It is necessary to follow up the study of the written material with observations and interviews, as described in the preceding chapter.

Once a probability tree diagram has been constructed, any potential system failures resulting from a single human failure will be obvious. For example, if an auxiliary operator has to reopen a critical valve, if no one checks his work, and if there are no other recovery factors in the system, that error alone could result in system failure. Such errors represent single channel failure modes. An advantage of the tree diagram is that these system-critical errors become obvious when the tree is drawn.

The probability tree diagrams also help to identify possible common-cause failure events. An example is the common-cause failure potentiality in the misadjustment of the test equipment used for calibration of sensors described in the discussion of Figure 5-1.

CHAPTER 6. UNAVAILABILITY

In safety analysis it is often necessary to incorporate HEPs in the estimation of component unavailabilities. This occurs when an analysis deals with maintenance or calibration errors that can render a component unavailable for some interval of time. In such cases the component's unavailability is a function of the probability of some human error event, the probability of recovery, and the average time that the component is in a failed condition before being restored.

This chapter presents some mathematical expressions used in unavailability calculations. These expressions can be used with the human performance models and HEPs in this handbook to estimate the unavailability of NPP systems and components due to human error. The topic of unavailability as it more generally relates to NPPs is treated in Chapter XI of the Fault Tree Handbook (Roberts et al, 1980).

Unavailability Equations

The availability of a system or equipment is the probability that it is operating or will operate satisfactorily if called upon. A common index of availability (A) is the ratio of mean uptime ($\overline{u}$) to the sum of mean uptime and mean downtime ($\overline{d}$); i.e.,

$$A = \frac{\overline{u}}{\overline{u} + \overline{d}}^{*} \qquad (6\text{-}1)$$

Unavailability (U) is:

$$U = 1 - A = 1 - \frac{\overline{u}}{\overline{u} + \overline{d}} = \frac{\overline{d}}{\overline{u} + \overline{d}} \qquad (6\text{-}2)$$

---

*In this chapter the symbols and notation used are different from those used in the rest of this handbook.

This expression for unavailability is appropriate when the probability of being in a failed state is independent of time, the steady-state condition, as in the examples given in this chapter. When time-dependent unavailability is of concern, more elaborate treatment is required as described in Vesely and Goldberg (1977). In this handbook, the time-dependent situation will be disregarded since the "steady-state" assumption is used most.

To evaluate the human error contribution to unavailability we use the same equations that are used for component and system failure occurrences (human errors simply being a particular cause of failure.)

Consider a human action that is performed periodically with an average time, $\overline{T}$, between actions. Assume that for any action there is a probability, p, that a human error will occur and not be recovered. Also assume that, given an error, the average downtime (the time that transpires before the error is corrected) is $\overline{d}$. Using the concepts underlying Equation 6-2 we have

$$U = \frac{p\overline{d}}{\overline{T}}, \qquad (6-3)$$

since $p\overline{d}$ is the average downtime and $\overline{T}$ is the average uptime plus average downtime (i.e., $\overline{u} + \overline{d}$).

To calculate U from Equation 6-3 we need p, $\overline{d}$, and $\overline{T}$. The probability p is obtainable from this handbook or other sources of HEPs. $\overline{T}$ and $\overline{d}$ are defined by the system. For convenience, p is sometimes divided into two factors, an error probability, E, and a nonrecovery probability, R:

$$p = ER, \qquad (6-4)$$

where E = the probability of committing the error per act, and

R = the probability of failing to recover the error at or about the time the error is committed. (For unavailability calculations, a delay of an hour or so for the potential recovery operation is ignored.)

$\overline{T}$, in Equation 6-3, is normally taken to be the time between tests on a component, and it is usually assumed that any error made on the first test will be discovered during the second test. Thus, if a valve is left in the wrong position after a monthly test, it is assumed this error will be discovered on the next monthly test, and $\overline{T}$ is 720 (the number of hours in an average month, 24 x 30). If full recovery of the original testing error cannot be assumed for the next test, $\overline{T}$ would be modified accordingly.

In Equation 6-3, $\overline{d}$ equals $\overline{T}$ if there is no checking between tests. If checking is done, $\overline{d}_T$, the total $\overline{d}$, is the sum of the $\overline{d}$ values for each time period between the first test and all subsequent checks and between the first test and the next test. For example, if there is a midmonth check between two monthly tests, $\overline{d}_T = \overline{d}_1 + \overline{d}_2$, where $\overline{d}_1$ is the time period between the first monthly test and (but not including) the midmonth check, and $\overline{d}_2$ is the time period between the first monthly test and (but not including) the second monthly test.

The general equation for calculating $\overline{d}_T$ is:

$$\overline{d}_T = h_1 + C_1h_2 + C_1C_2h_3 + \ldots + C_1C_2\ldots C_m\overline{T} \qquad (6\text{-}5)$$

where h_1, h_2, and h_3 are, respectively, the number of hours between the first test and the first, second, and third check; $\overline{T}$ is the number of hours between successive tests; and C_1, C_2, and C_m are, respectively, the probabilities of non-detection at the first, second, and last check performed between the two tests.

Recall that Equation 6-3 applies to steady-state conditions and should not be used if there is time dependence. Unavailabilities resulting from human actions can be used in system safety or system reliability analyses in the same way that hardware unavailabilities are used. If these human-caused unavailabilities are combined with hardware unavailabilities, we should check whether the hardware unavailabilities already include human error contributions to avoid counting them twice. For situations not covered by the equations given here, other unavailability expressions must be developed using the usual reliability theory approaches, as described in Barlow and Proschan (1965).

In treating unavailabilities resulting from human actions the same way one treats hardware-caused unavailabilities, a conservative approach is usually taken. It is generally assumed that the detection of an error does not influence the probability of future errors being committed or detected. Actually, when an error is found, extra precautions are often taken, and the error probability, p, for subsequent tests may be reduced for some time after the error is detected.

Applications of the Unavailability Equations

To illustrate the use of the above equations two examples are given.

Example No. 1 - Unavailability of a Diesel Generator

Assume that we wish to estimate the unavailability of a back-up diesel generator. (In this example we address only that unavailability due to human error.) Each Monday a technician takes the diesel offline and tests it. Thus, the time between tests is $\overline{T} = 168$ hours. For this example, assume that the probability that he will forget to place the diesel back online is estimated as $E = .01$.

If the technician makes the error, the diesel will be unavailable until the next time it is tested or until the error is detected during some intermediate inspection. Suppose that immediately after the technician performs his task, an inspector, as part of a written procedure that includes other tasks, checks that the diesel is back online. This inspector could fail to notice that the diesel is offline. Assume that the probability of nonrecovery of the error is R = .1. Thus p = ER = .01 x .1 = .001. This is the probability of an unrecovered error at the time of the test. The short time between the actions of the technician and those of the inspector is regarded as negligible. If no further inspection is made, and if there are no other recovery factors, the diesel will be unavailable for one week; i.e., until the next test. Thus, $\overline{d}$ = 168 hours. The unavailability is calculated from Equation 6-3 as:

$$U = \frac{.01 \times .1 \times 168}{168} = .001$$

Note that, in this particular example, $\overline{T}$ and $\overline{d}$ are the same because no checking occurs between the tests. If checking does occur between the tests, $\overline{d}_T$ is calculated using Equation 6-5. If a check occurs at the beginning of the fifth day (i.e., 96 hours after the test), h_1 will equal 96 hours. Since there is only one check between the tests, h_2 equals $\overline{T}$ (168 hours). Assume that at the beginning of h_2 a special check is scheduled to ascertain if the diesel is available. If the probability of the checking error C_1 is .05,

$$\overline{d}_T = 96 + (.05 \times 168) = 104.4 \text{ hours}$$

and using Equation 6-3,

$$U = .01 \times .1 \times \frac{104.4}{168} = .0006$$

Example No. 2 - Unavailability of a Primary Source of Emergency Coolant

Assume a quarterly test of the primary water supply in the Emergency Core Cooling System (ECCS) at a plant, and that the only human error of consequence for the unavailability of this coolant is that of leaving the main isolation valve in the wrong position after the test. For this error assume $E = .01$ and $R = .1$ as above. Assume there are two scheduled monthly checks between tests, with $C_1 = C_2 = .05$. In this case, h_1 is 720 hours, h_2 is 1440 hours, and $h_3 = \overline{T} = 2160$ hours. Using Equation 6-5,

$$\overline{d}_T = 720 + (.05 \times 1440) + (.05 \times .05 \times 2160) = 797.4 \text{ hours}$$

and using Equation 6-3,

$$U = .01 \times .1 \times \frac{797.4}{2160} = .0004$$

The importance of the two checks between the tests can be seen by calculating U without these tests. In this case, an unrecovered error (ER) at the time of the first test will remain undiscovered until the next test. The probability that the main isolation valve will be unavailable is simply the probability of the unrecovered error, and $U = p = ER = .01 \times .1 = .001$. Thus, the two checks between the tests reduce the unavailability of the primary source of emergency coolant by a factor of 2.5 (i.e., .001/.0004).

PART III. HUMAN PERFORMANCE MODELS AND ESTIMATED HEFs

PART III. HUMAN PERFORMANCE MODELS AND ESTIMATED HEPs

Chapters 7 through 18 present models and estimated HEPs for the quantification of human performance in NPPs. These represent our hypotheses about the performance of NPP personnel under a variety of tasks and conditions. Because of the lack of objective data on human performance under controlled conditions in NPPs or NPP simulators, our models and estimated HEPs are based in part on:

- Our experience in human reliability analyses of NPP and weapon operations
- Our background as experimental psychologists
- Recorded incidents (such as those reported in the NRC Licensee Event Reports)
- An extensive literature search
- Interviews with and observations of NPP personnel
- The experience of our colleagues and other workers in the field of human performance

Whenever possible, the estimated probabilities of human error are based on relevant data from laboratory studies or more applied settings. When no documented data were available, judgments were used to estimate HEPs. Supporting data for the estimated HEPs are discussed in Chapter 19. Each HEP is listed with its estimated uncertainty bounds; the rationale for these bounds is discussed in Chapter 16. We estimate that the bounds correspond to the 5th and 95th percentiles of the conjectured distribution of an HEP.

The information obtained from these sources was evaluated in the context of accepted theories of human learning and performance (e.g., Stevens, 1951; McGeoch, 1942; McGeoch and Irion, 1952; Berelson and Steiner, 1964; Chapanis et al, 1949; McCormick, 1975; and Woodworth, 1938). As we have said before,

" ... there is still a very large gap between the needs for a central data bank of human performance data and what is available" (Swain, 1969b). Until such a bank can be developed, human reliability analyses will depend to a large extent on judgment. We make no apology for the fact that much of the quantification and models in this handbook represent our judgment based on our expertise; such is the state of the technology today. This handbook is a start -- we hope that others will verify our speculations.

In developing the estimated HEPs and the models, we tried to consider both typical and atypical modes of operations. We have the most confidence in applications of the models and HEPs to typical rule-based behavior in which the analyst estimates HEPs for routine operations. Applications to unusual modes of operation are highly speculative; these are discussed in Chapters 16 and 17.

At this early stage in the technology of human reliability analysis it is not possible to model all the effects of PSFs either in general or in their applications to a specific NPP. This limitation must be considered in applying the models and numerical results to any given NPP or group of NPPs. Assessment errors will be reduced to the extent that the PSFs are evaluated correctly.

Our approach in model development was to use the above-mentioned sources in formulating psychologically sound human performance statements based on the operational environment, the tasks of the operators, and the relevant PSFs. In consultation with statisticians knowledgeable about NPP operations, we developed simple, standard mathematical expressions that reasonably describe human performance statements. Advantages to using standard statistical distributions for these descriptions include the ready availability of tabled values and the added calculational convenience. For example, it was apparent in developing the curve of recovery of efficiency for walk-around inspections that the cumulative normal distribution reasonably approximated the recovery function. Therefore, this distribution was specified as the recovery function for walk-around inspections.

Another example is illustrated by experimental evidence (the vigilance literature cited in Chapters 3 and 17) and NPP experience which indicate that, in evaluating the probability that an operator will detect an unannunciated display of an out-of-tolerance condition, the highest probability of detection occurs during the initial audit. Operators characteristically scan most actively during this period. After the initial audit a combination of the effects of fatigue and expectancy causes the probability of detection to decline rapidly and then level off at some low level. Our statistical consultants showed us that an exponential function was a reasonable fit for this performance curve. We therefore specified the exponential curve as descriptive of this performance rather than specifying some other mathematical function that might fit closer but would be less convenient to use.

Our approach to the mathematical modeling of human performance is more inductive than that used by classical modelers. We first induce the statements of human performance from our experience and the available data; then we select mathematical expressions that fit the statements. Thus, our basic premise is that the psychology must be sound; the mathematical expresssions are secondary.

CHAPTER 7. DEPENDENCE

A major problem in human reliability analysis is the determination of how the probability of failure or success on one task may be related to failure or success on some other task. Two events (e.g., failure on Task "A" and failure on Task "B") are independent if the conditional probability of one event is the same whether or not the other event has occurred. For example, independence would apply to the case in which the probability of success on Task "B" were the same regardless of success or failure on Task "A." If events are not independent, they are dependent, and the nature of the dependence must be considered in the human reliability analysis.

The best method for assessing dependence is to determine the conditional probabilities from actual data. For example, if Tasks "A" and "B" have been performed under the applicable conditions a large number of times, the probability of failure on Task "B" can be calculated separately for the situations in which there is failure on Task "A" and in which there is success on Task "A," and the conditional probabilities can be determined.

A second method, used when objective data are not available, is to make judgmental assessments of the conditional probabilities on the basis of the nature of the tasks and their interrelationships. We often use this approach.

A third approach, developed for this handbook, uses a dependence model which, along with estimates of the appropriate level of dependence, serves as an aid in the estimation of conditional probabilities. This model is based primarily on human factors considerations.

Another approach to modeling dependence, based primarily on probabilistic considerations, has been developed by R. G. Easterling and is presented in the appendix to this chapter.

Two Types of Dependence

Two types of dependence are recognized -- direct and common cause. Direct dependence exists when the outcome of one task directly affects the outcome of a second task. This is the case when the failure or success of one task changes the environment or other factors so as to change the probability of failure on another task. Some examples of direct dependence between Tasks "A" and "B" are:

(1) Failure on Task "A" causes an auditory signal that results in more careful performance of Task "B."

(2) Failure on Task "A" causes extreme anxiety with a resultant increase in probability of failure on Task "B."

(3) Failure on Task "A" causes Task "B" to be more difficult with an associated increase in probability of failure.

The second type of dependence exists when the performance of two or more tasks is related to some common influence, or common cause. The common influence may consist of the person's attitude, perceptual set (i.e., what he expects to see or happen), emotional state, skill level, or it may consist of situational variables and many other factors that shape performance.

An example of a common cause that influences human performance is the case in which some set of tasks must be performed under high stress. For most people, a high level of stress tends to increase the probability of error for all of the tasks to be performed. Although the individual tasks might be independent, there would be a dependence between the quality of performance of each task and the common influence of stress. In addition to raising the basic probability of error for each task, a high level of stress may compound the interaction among the tasks.

As mentioned in Chapter 1, this handbook takes a scenario-oriented approach to the prediction of human performance. Each scenario includes the

PSFs judged to be the most important in estimating the HEPs for any given task. The common cause influences are represented in these PSFs. From our point of view, the distinction between common cause dependence and direct dependence is not very useful. In one sense, when we deal with human performance, almost all dependence is common cause. If I make an error on Task "A," this performance does not always directly influence my performance on Task "B." Except for essentially automatic human reactions, or some physical dependence of the task objects, the performance on Task "B" is mediated by my mental/emotional reaction to my perception or knowledge of my failure on Task "A." It is not often similar to the direct cause and effect relationship that exists when a break in one pipe carrying caustic material can cause a nearby pipe to fail. In practical human reliability analysis we don't attempt to differentiate between the two types of dependence; we determine the appropriate PSFs and their overall effects, including the appropriate weightings for direct dependence between human task performances. The cases in which a common cause may affect performance are allowed for in the assessments of direct dependence.

Characteristics of Dependence

The degree of dependence among human actions, or between system events and the human reaction to them, ranges along a continuum from complete independence (zero dependence) to complete dependence. Theoretically, dependence can be either negative or positive.

Negative dependence implies a negative relationship between events; that is, failure on the first task reduces the probability of failure on the second task, or vice versa. Most of us can think of everyday experiences in which, having erred in some task, we take extra care and have a higher probability of success on the second task. Conversely, success on the first task might reduce our probability of success on a second task through overconfidence. There are also instances of negative dependence on task performance between people in industrial situations, as when a rivalry exists. In this chapter we primarily address the effects of dependence among tasks performed by one person.

Positive dependence implies a positive relationship between events; e.g., failure on the first task increases the probability of failure on the second task. In our work we address positive dependence only, as the usual NPP atmosphere is one of cooperation, and cases of negative dependence are relatively rare.

Considering only positive dependence, then, we must judge the degree of dependence among tasks or events when performing a human reliability analysis. If one has the data, the best approach is to calculate directly the conditional probabilities of the actions under consideration and to incorporate the appropriate degree of dependence in the error probabilities. For example, assume that Person A will perform a task and Person B will check A's work. If data are available, the probability of Person A failing ($\Pr[F_A]$), Person B failing given that Person A did not fail ($\Pr[F_B|S_A]$), and Person B failing given that Person A failed ($\Pr[F_B|F_A]$), can all be determined. These probabilities can then be used to compute the appropriate joint probabilities. Such composite data are usually not available. One may find data on how accurately Person A performs his task but none on how accurately Person B checks the task. However, it may be known that Person B's accuracy is about the same as that of Person A when performing the same work. In such a case, one approach is to use the estimated probability of failure for Person A and apply it to Person B, with some modifying factor to account for the judged level of dependence. If there is no dependence and if their error probabilities are equal, the joint probability that Person A will make an error and that Person B will fail to catch the error is the square of the estimated error probability for Person A. Such a state of affairs is most unusual.

At the opposite extreme of the continuum (complete dependence), one might judge that if Person A makes an error, so will Person B. That is, Person B will not detect the error made by Person A. In this case, the estimated joint probability that Person A and Person B will make errors is the estimated probability for Person A alone, since the conditional error probability for Person B is 1.0 under the assumption of complete dependence. Complete dependence is not as unusual as zero dependence; however, for the example stated, it would not be likely. The more usual state of affairs is a level of dependence somewhere between zero and complete dependence.

Levels of Dependence

Dependence is a continuum, and it is necessary to judge the appropriate level between any pair of tasks. This may be difficult, and some simplification may have to to be made. The approach taken here is to reduce the continuum of conditional probability to a small number of discrete points. For most human reliability problems we use just five points: The two end points of zero dependence (ZD) and complete dependence (CD), plus three points in between. We call these intermediate points low dependence (LD), moderate dependence (MD), and high dependence (HD). The equations for the five points are presented later.

Figure 7-1 shows the relative positions of the five discrete points in this model. The rationale for the assignment of these relative positions is given later. Use of the values between HD and CD would result in little change in the estimated joint HEPs. Therefore, no discrete points are assigned in this area.

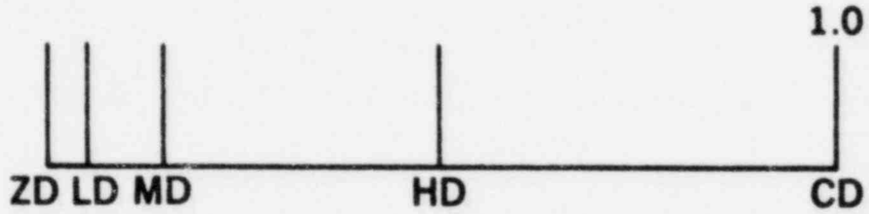


Figure 7-1. Continuum of Dependence Represented by Five Discrete Points

It is relatively easy to decide whether ZD or CD is appropriate. If neither extreme is appropriate, one has to assign one of the three intermediate levels of dependence. There are no set rules for differentiating among the three intermediate levels, LD, MD, and HD. Although this judgment depends on the experience of the analyst, we will present some general statements and examples to illustrate the distinctions among these three intermediate levels of dependence. We wish to stress that if data exist regarding the level or the effects of dependence for any situation, these should be used in lieu of the dependence model.

Complete-Failure Path and Complete-Success Path Dependence

For the three intermediate levels of dependence, models are provided for (1) the complete-failure path dependence (i.e., dependence among tasks that are all performed incorrectly), (2) the complete-success path dependence (i.e., dependence among tasks that are all performed correctly), and (3) all other paths. In practical human reliability analyses, the first two paths are usually the ones of interest. That is, the analyst is generally interested either in the probability that a person performs at least one of the related tasks correctly (i.e., he avoids the complete-failure path) or that he performs all related tasks correctly (i.e., he follows the complete-success path).

Figure 7-2 shows a three-task probability tree diagram consisting of the success and failure limbs for Tasks "A," "B," and "C." (The symbology from Table 5-1, p 5-6 is used in this chapter.) On the complete-success path in the tree, a nonzero level of dependence means that the probability of success on succeeding tasks, given success on all preceding tasks, is higher than the basic human success probability (BHSP) (i.e., what the success probability

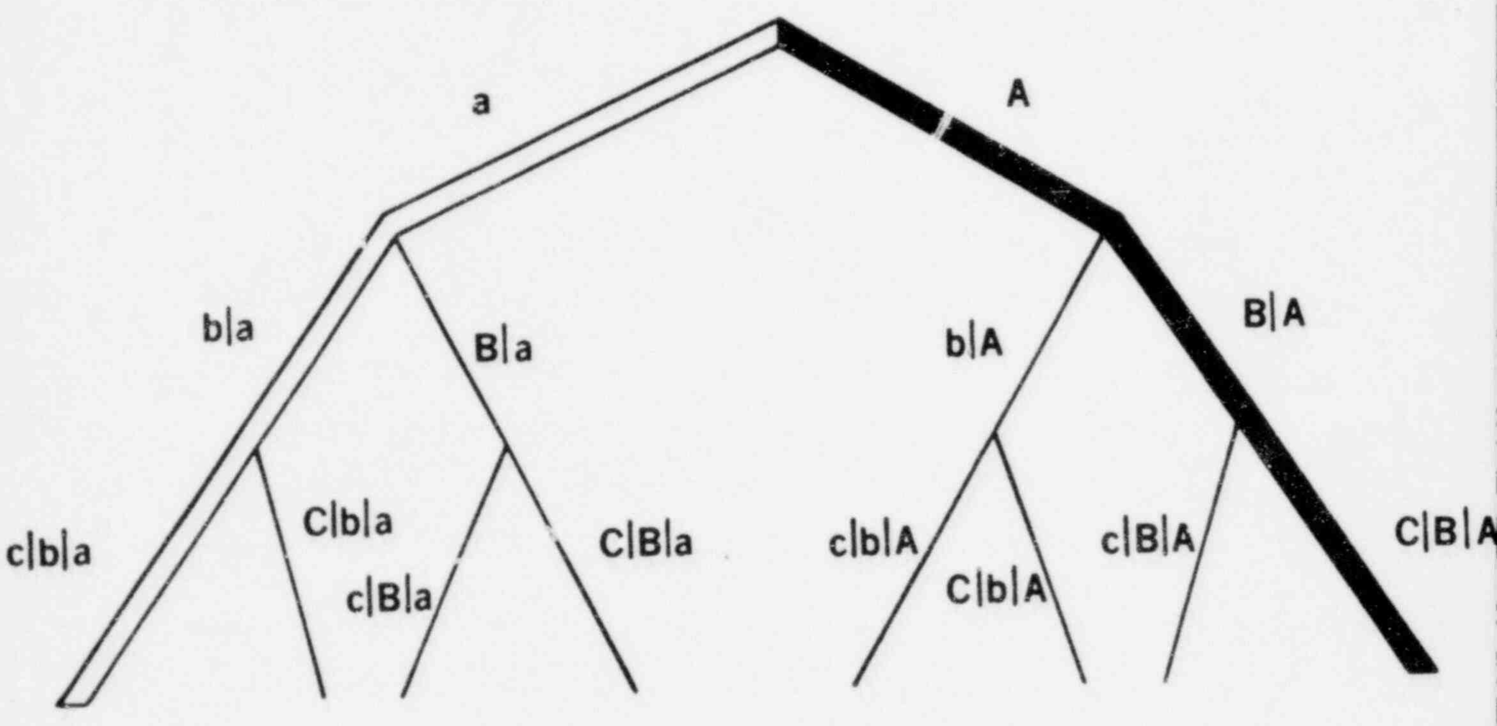


Figure 7-2. Probability Tree Diagram Showing Complete-Success Path (double line) and Complete-Failure Path (heavy solid line)

would be in the absence of dependence). On the complete-failure path in the tree, a nonzero level of dependence means that the probability of failure on succeeding tasks, given failure on all preceding tasks, is higher than the BHEP (i.e., what the error probability would be in the absence of dependence).

Considering only Tasks "A" and "B" in Figure 7-2, the conditional probability of success or failure of Task "B" varies according to the influence of Task "A." Consider the complete-failure path and assume a nonzero level of dependence between these two tasks. If Task "A" is failed, the probability of also failing Task "B" will be higher than its BHEP (i.e., B|A will be higher that B). Conversely, the probability of successful performance of Task "B," given failure on Task "A" (i.e., b|A), will be lower than b, the BHSP.

For the complete-success path, still assuming a nonzero level of dependence, the probability of successful performance of Task "B," given successful performance of Task "A" (i.e., b|a), will be higher that the BHSP, b. Conversely, the probability of failure of Task "B," given that Task "A" is performed correctly (i.e., B|a), will be lower than the BHEP, B. (The reader should keep in mind the fact that we consider only positive dependence.)

When estimating the conditional probability of success or failure of Task "C" in Figure 7-2, one must decide what level of dependence is appropriate between this task and the preceding task, Task "B." The level may or may not be the same as the level between Tasks "A" and "B." (As noted in the following section, the conditioning influence of Task "A" on Task "C" is ignored.)

Dependence in Other Paths

For a human reliability problem consisting of only two tasks, the conditional probabilities of the two limbs other than the complete-success and the complete-failure paths are predetermined because at each branching the conditional probabilities for the limbs must sum to 1.0. Thus, in Figure 7-2, b|a + B|a = 1.0, and b|A + B|A = 1.0. For any two-task tree, if we have estimated b|a and B|A, the probabilities for the other two limbs can be obtained by subtraction.

In the case of an analysis involving more than two tasks, the conditional probabilities of subsequent limbs are derived by attributing all the effects of dependence to the immediately preceding task (or limb). For example, in Figure 7-2, the conditional probability C|B|a is assumed to be the same as that of C|B|A, and the conditional probability of c|b|A is assumed to be the same as that of c|b|a. In both of these cases the influence of Task "A" on Task "C" is ignored. Although we recognize that this assumption is not entirely valid, our rationale for simplification is that the immediately preceding task is generally the prime factor influencing the success or failure of the task in question, given some nonzero level of dependence between the two tasks. It is possible to derive a model that assigns different weights to the conditioning factors of, say, C|b|a and C|b|A, but considering the typical uncertainty in estimates of the basic success or failure probabilities, such exactitude is unwarranted. Such refinement is best postponed until the basic human performance data warrant such improvement. If we do not believe that C|b|a = C|b|A, we estimate conditioning effects directly, as in the examples beginning on p 7-19.

Relationship of Dependence Model to an Earlier Model

The dependence model in this chapter replaces an earlier model developed in "Human Reliability Analysis," Section 6.1, in Appendix III of WASH-1400, and in Swain and Guttmann (1975). The earlier model used the geometric mean of the joint probabilities of error under ZD and CD to estimate the joint probability of failure of two or more human activities when a moderate level of dependence was assumed. Under the assumption of a high level of dependence, the joint probability of failure of two or more human activities was calculated by using the geometric mean of the joint probabilities of error under MD and CD. The present model will yield somewhat more conservative (i.e., higher) failure estimates than will the geometric mean model. The present model is more easily understood in a practical sense, and it has greater flexibility since there are three intermediate levels of dependence. In the earlier model we had only ZD, MD, HD, and CD; in the present model we have added LD.

Psychological Considerations for Levels of Dependence

This section discusses the psychological meaning of each of the five levels of dependence and presents some examples. Guidelines are provided to aid in the judgment of which level of dependence is most appropriate for typical situations.

As described in Chapter 2, there are five categories of error: errors of omission, errors of commission, sequential errors, extraneous acts, and time errors. In performing a human reliability analysis, different levels of dependence may be used in estimating failure probabilities for different categories or error. For example, in the model of valving operations in Chapter 13, ZD is assumed for several errors of commission related to the manipulation of a sequence of valves. Higher levels of dependence are often assumed for errors of omission.

General Guidelines in Assessing level of Dependence

While there are no set rules to apply in assessing which level of dependence to use, there are a few general guidelines:

(1) Evaluate the influence of the failure or success of the immediately preceding task on the task of interest. The level of dependence may not remain constant throughout a series of activities.

(2) Use the higher of two levels of dependence when there is doubt about which of them is the more appropriate. This will result in a more conservative assessment of the failure probability for the complete-failure path and a more optimistic assessment of the success probability for the complete-success path.

(3) Evaluate the spatial and time relationships among all events. Dependence between any two events increases as the events occur closer in space and time. For example, displays or controls that are physically close to each other or that must be manipulated at about the same time have a higher level of dependence than items that are widely separated either spatially or as to the time of their manipulations.

(4) Evaluate the functional relationships among events. Dependence between any two events increases with their functional relatedness. For example, events within a subsystem have a higher level of dependence with respect to each other than to events in some other subsystem. (Further discussion of the concept of functional relatedness is presented in the next topic heading.)

(5) Evaluate the similarities among NPP personnel with respect to all relevant factors. Dependence among personnel increases with similarity in status, training, responsibility, and many social and psychological factors.

(6) Reevaluate the level of dependence assumed if the joint probability of failure for two tasks is as low as 10^{-6} in a case in which one person is checking the other. Likewise, if the joint probability of failure of two related tasks performed by one person is as low as 10^{-5}, reevaluate the level of dependence assumed. Such low end-failure probabilities in these cases should be suspect.

Functional Relatedness Among Tasks

The concept of functional relatedness is often central to judging the level of dependence among events or items within a task or among tasks themselves. In the performance models dealing with displays, valves, controls, etc, the analyst must evaluate the functional relationships among these items. Functional relatedness is defined for an operator in terms of his training and experience and in the way the work operations or written procedures are designed. The highest level of relationship exists when two or more events or items are, in effect, regarded as a single unit by an operator. For example, in some NPPs, certain pairs of MOVs represent two channels of the same function, and their controlling switches in the control room are operated simultaneously. The operator characteristically uses both hands, one for each switch. For these situations, the switches can be considered completely dependent with regard to errors of omission. That is, if the operator remembers to manipulate one of the two switches, he will remember the other. Whether or not this highest level of functional relatedness also applies to errors of commission will depend on the design of the switches. If, for example, both switches are rotated clockwise to open the MOVs, the assumption of CD for errors of commission would normally be appropriate. If the switches operate as mirror images (one is rotated clockwise and the other is rotated counterclockwise to open their related MOVs), the assumption of CD will not likely be appropriate, and the analyst will have to estimate the HEPs for commission on some other basis.

In the above example, the assumption of CD for errors of omission might be appropriate even if the two switches are separated such that simultaneous operation by an operator is not possible. The key here is the time relationship and the typical mode of operation. If the operator typically manipulates one switch and then immediately manipulates the other, this association might be judged strong enough to warrant the assumption of CD for errors of omission.

Other items of equipment may not have the same function in an NPP, but, because of their location or other association, they may be regarded as one unit, at least for errors of omission. For example, if several manual valves are located in a small group and if the written procedures treat them as a group (in a single written step), the assumption of CD will normally be appropriate. An exception is the case in which some of them are supposed to be open and some closed, breaking the perceptual pattern of a "single unit."

Displays may have different levels of functional relatedness. For example, temperature and pressure gauges for the reactor vessel are obviously related, yet this does not necessarily mean that an operator who checks one display will invariably check the other. In Chapters 9 and 11 special rules are formulated for the assessment of the level of dependence among displays.

In the handbook, the term one item of equipment refers to an individual item of equipment (e.g., a display, control, valve, etc) or some group of items that are completely dependent with regard to errors of omission. The level of functional relatedness for errors of commission must always be evaluated separately.

Regrettably, we can offer no detailed guidelines for judging the extent of the functional relationship between events or items of equipment. It is the operator's perception of what is functionally related that must be evaluated. In our own analyses we tend to be conservative as we wish to avoid underestimating the influence of operator error.

Zero Dependence

Zero dependence (ZD) applies to the case in which the quality of performance, including nonperformance, of one activity has no effect on the performance of subsequent activities. Although 100% independence is unusual for human events, occasionally the level of dependence is so slight that we assume ZD for purposes of analysis.

An example of a situation in which ZD is usually assumed between tasks for errors of commission is that of an operator who has to check-read several instruments under normal operating conditions as part of his periodic scanning of displays during a shift. If the probability of a check-reading error on some instrument is estimated as .003 per act, and there is a similar instrument displaying different information elsewhere on the control panel that must also be check-read, the same .003 error probability should be used for each.

On the other hand, ZD is not a valid assumption for errors of commission if the characteristics of the work situation include all of the following: (1) the meters are located side by side (2) the pointers on meters are "lined up" under normal operating conditions, and (3) the stated operating policy is to check-read both meters simultaneously. In such a case, we would probably assume complete or at least high dependence between the check-readings.

For the same example, the level of dependence appropriate for errors of omission should be assessed separately. For example, if the two check-reading requirements are separated in time so that they represent unrelated tasks to the operator, ZD would be assumed. However, if the two check-readings are linked (if they must be read at the same time), the assumption of complete dependence might be appropriate. If the operator forgets to check one, he will forget to check the other.

We usually assume ZD when estimating the error probabilities for carrying out individual steps in a written procedure. This applies to errors of commission and omission, given that the operator has initiated the task and is using the written instructions.

Low Dependence

Low dependence (LD) represents a level of dependence that is greater than ZD, but is not very far up the dependence continuum. It is a convenient assumption to make when the dependence between actions is clearly greater than zero but not much greater. For the complete-failure path (see Figure 7-2, p 7-8), the use of LD rather than ZD will avoid undue optimism in reliability estimates and will also compensate for the inability to foresee various sources of human interaction, such as when a human reliability analysis is done without a detailed task analysis. For the complete-success path, the use of LD rather than ZD makes no material difference in the calculation of the joint probability of success.

Moderate Dependence

Moderate dependence (MD) represents a level of dependence between LD and HD. We use this level when an intermediate level of dependence is appropriate, but neither the low nor the high level suits the case at hand.

High Dependence

High dependence (HD) represents a level of dependence that is approximately midway between ZD and CD on the continuum. It is a convenient assumption to make when the dependence between two actions is not complete but is definitely towards the high end of the dependence continuum. HD means that the performance on one task very substantially affects the performance of a subsequent task. In some cases, the use of HD rather CD will avoid undue pessimism. For example, if the dependence levels apply to the complete-failure path, the assumption of CD means that an error on Task 2 is inevitable, given an error on Task 1. The assumption of HD reduces the estimated conditional probability of error on Task 2 to some value lower than 1.0.

Complete Dependence

Complete dependence (CD) between the actions of two people working on a job is unusual, although not as unusual as ZD. CD between two actions performed by the same person is more common. For example, in certain NPP control rooms, switches are often paired, and the written procedures treat the pair as a unit. Thus, the procedures for a given situation in one PWR call for the operator to, "Open MOV-860A and B, suction to the low head SI pumps from the containment sump." (This is part of the procedure for changing from the injection to the recirculation mode after a LOCA.) Typically, the operator reaches for both switches and operates them as a unit. If he fails to operate either, he is almost certain to fail to operate the other. Conversely, if he operates either, he is almost certain to operate the other. The primary error here is one of omission. If he forgets one, he will forget the other in almost 100% of the cases. Therefore, CD is the appropriate level of dependence. On the other hand, if these two switches are located such that simultaneous manipulation cannot be achieved with one motion, a lesser degree of dependence should be assessed.

Errors of omission provide a common example of CD. If an operator fails to initiate a task (an error of omission), all the steps in that task will also be omitted. The converse does not necessarily hold; the fact that an operator initiates a task does not guarantee that all the steps will be performed.

An Example Illustrating Several Levels of Dependence

Several levels of dependence can be illustrated by the tasks involved in preparing a Federal Income Tax form. If a person takes his records to a firm specializing in the preparation of income tax forms, that firm depends completely on the data provided them. If the customer presents an incorrect but believable record of expenses, the firm will not be able to check its validity. Thus, the customer's error (or any deliberate fabrication) will go unchecked because the firm's calculations depend completely on the information supplied.

In the above case, if the customer also prepared his own tax return separately from that prepared for him by the income tax firm, we could assume ZD between those errors of the customer and of the firm that are due to arithmetic and interpretation.

Other levels of dependence can be illustrated by assuming that the person prepares his income tax return without outside help. If he has great confidence in his arithmetic, etc, he may merely check over his tax return after completion. Assume that he will review his arithmetic by repeating the mental calculations. If he calculates the same way he did the first time (e.g., in adding a column of figures, he starts at the top and works down), we assign a high level of dependence between his original work and his immediate checking of it. On the other hand, if he follows a different procedure for his checking task (e.g., adding a column upwards this time), a low level of dependence is more appropriate. We cannot use ZD because there are still several possibilities for common cause failures. For example, when he did his work originally, he wrote a 3 that looks like an 8. If he read it as an 8 when he added it the first time, he is likely to do so again. In the first case above, given an error on the original trial, the assumption of HD means that there is a relatively low probability of his catching any error in arithmetic. In the second case, LD means that the probability of his detecting an error is considerably higher.

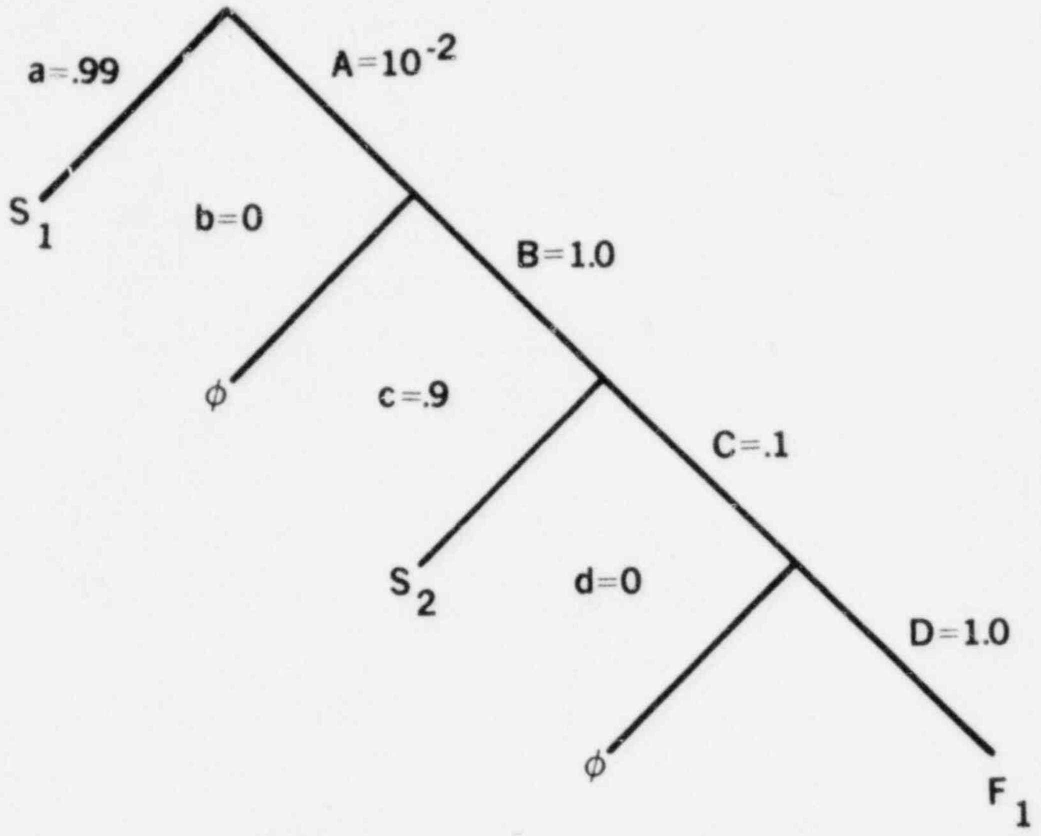
To continue with the income tax example, assume that the person is aware of the likelihood of repeating his errors if he makes an immediate check, and decides to wait 24 hours and then repeat his work without reference to the previous day's work. Within 24 hours he will probably have forgotten the arithmetic calculations, so that the second day's work will be essentially independent of the first day's work, and ZD may be assumed between the two trials. Other mental processes are less likely to be independent; i.e., errors in logic or interpretation are much more prone to be repeated. If some instruction in the tax form is ambiguous, the interpretation made on the first reading is likely to be made on subsequent readings, possibly because of some predisposition of the reader. Thus, for the case of interpretation of ambiguous instructions, either HD or CD would be appropriate.

Direct Estimation of Conditional Probabilities of Error

The following example illustrates a case in which we did not use the dependence model, but estimated conditional probabilities of error directly. The example also illustrates the importance of recovery factors in a situation involving CD.

In this situation, a technician is checking the calibration of a series of setpoints consisting of three comparators. To do this, he must first set up some test equipment, and he could make an error in this initial setup. For example, he could select the wrong decade resistance, set up the wrong scale on the decade, or make some other errors in the test setup. Unless corrected, such an error will result in miscalibration of all three comparators.

The problem was evaluated in WASH-1400, p II-101. Figure 7-3 presents the event tree diagram for this task. In this evaluation, a probability of 10^{-2} was estimated for the common cause failure of a miscalibration due to faulty setup. This estimate was modified by recovery factors as follows: it



A=FAILURE TO SET UP TEST EQUIPMENT PROPERLY

B=FAILURE TO DETECT MISCALIBRATION FOR FIRST SETPOINT

C=FAILURE TO DETECT MISCALIBRATION FOR SECOND SETPOINT

D=FAILURE TO DETECT MISCALIBRATION FOR THIRD SETPOINT

Figure 7-3. Probability Tree Diagram of Hypothetical Calibration Tasks was reasoned that when the technician discovered that the calibration of the first setpoint had to be changed, he would change it. It was further reasoned that when he found that the second setpoint also had to be changed, 90% of the time he would be suspicious, would recheck his test setup and discover his error. Ten percent of the time he would not be suspicious, and, given that he had this unsuspicious nature, it was judged that the conditional probability of the third error (i.e., failing to become suspicious when he has to recalibrate the third setpoint) was 1.0. That is, CD was assumed between the last two tasks. Thus, the joint probability of error in calibrating the three setpoints was $.01 \times 1.0 \times .1 \times 1.0 = 10^{-3}$.

Note that the above example assumes CD between the setup task and calibration of the first setpoint. CD is also assumed between the tasks of calibrating the second and third setpoints. However, a different level of dependence is assessed between the tasks of calibrating the first and second setpoints. This is a unique situation in which the interaction between two tasks is not covered adequately by the dependence model. The .9 probability that the technician will be alerted by the misalignment of the second setpoint is based on judgment, not on the dependence model. The analyst should always use the estimate in which he has the highest level of confidence. In this example, we had most confidence in the use of a conditional probability of .1 to represent the dependence between errors on Setpoint 1 and Setpoint 2.

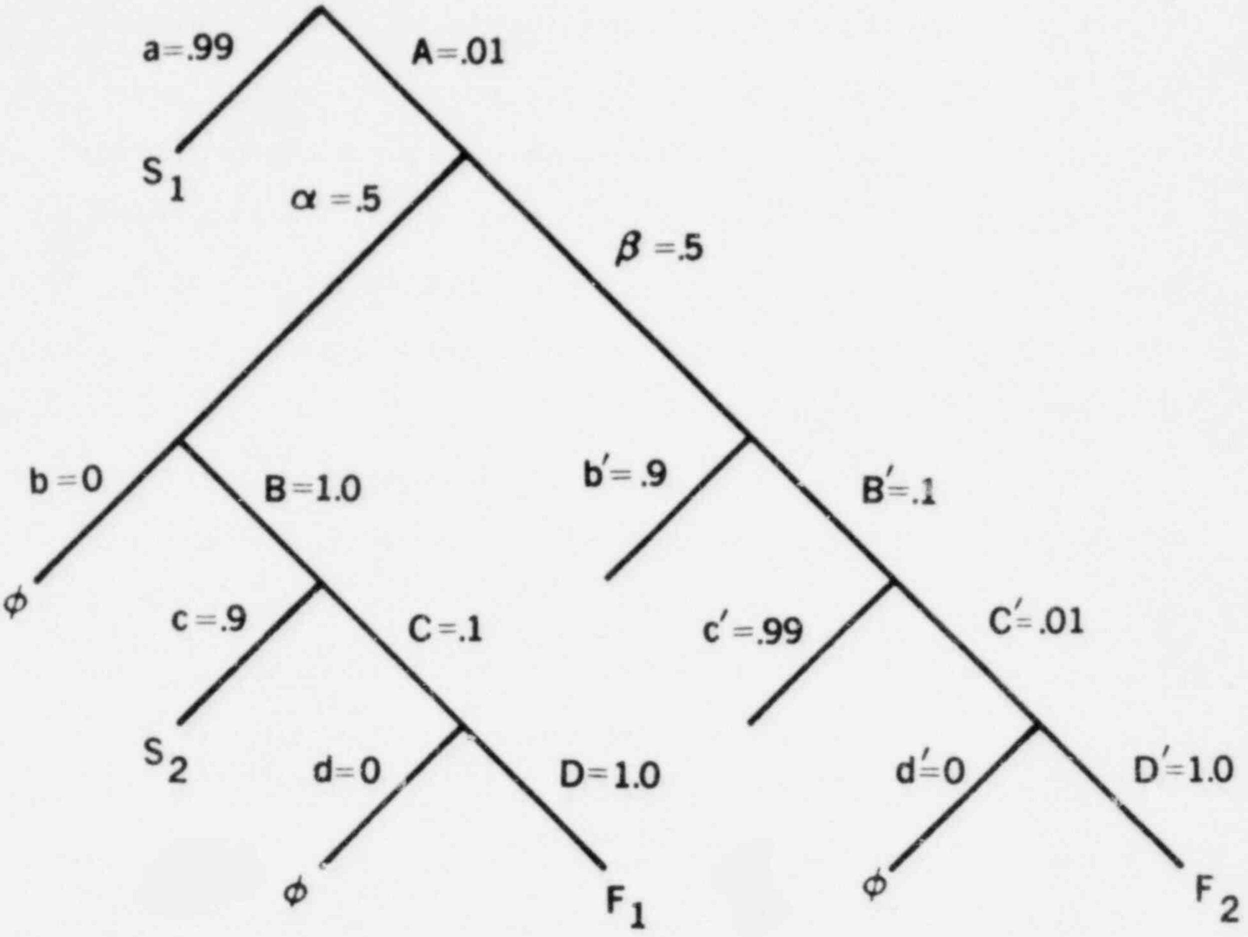
For illustrative purposes, we have evaluated this situation in more detail by considering both small and large miscalibrations. We present the analysis in event-tree form to demonstrate this graphic task analysis technique. We define a large change as one that is not normally expected, while a small change is one that can be expected to occur occasionally because of variation in equipment and other conditions.

In the event tree, Figure 7-4, Task "A" represents the test equipment setup. Alpha and beta refer, respectively, to the probability of small or large miscalibrations resulting from A, the incorrect setup of the test equipment. The other letters refer to the calibration of the three setpoints, with plain letters used on the alpha side of the tree, and prime letters on the beta side of the tree. As noted on p 5-7, we usually do not employ the complete notation for the conditional probabilities of events. That is, we do not write $\alpha|A$ or $\beta|A$; we write α or β, with the "given A" understood.

In the analysis, the only error of interest was that of the miscalibration of all three setpoints; that is, the complete-failure paths ending in F_1 and F_2, the two joint probabilitites of failure. Path a was designated as a success path (S_1) because, given that the technician set up his test equipment correctly, there were no other common cause probabilities of interest. Even if we assume that the BHEP of miscalibrating any given setpoint is as high as 10^{-2}, the probability of missetting all three is 10^{-6}, a negligibly small number. This example illustrates the usual case that an event tree need not depict all possible system outcomes.

Given A, we estimated that half the consequent miscalibrations would result in a small change, and half in a large change. (In a real analysis, people other than the human reliability analysts would provide these estimates.)

Considering the alpha side of the tree, B = 1.0 means that 100% of the time the technician will make the small calibration change and then proceed to Task "C," calibration of the second setpoint. $C = 10^{-1}$ means that 10% of the time he will fail to become suspicious when the second set point also requires adjustment. c, then, is .9, which means that 9 times in 10 the technician



A- FAILURE TO SET UP TEST EQUIPMENT CORRECTLY

α- SMALL MISCALIBRATION OF TEST EQUIPMENT

B- FOR A SMALL MISCALIBRATION, FAILURE TO DETECT MISCALIBRATION FOR FIRST SETPOINT

C- FOR A SMALL MISCALIBRATION, FAILURE TO DETECT MISCALIBRATION FOR SECOND SETPOINT

D- FOR A SMALL MISCALIBRATION, FAILURE TO DETECT MISCALIBRATION FOR THIRD SETPOINT

β- LARGE MISCALIBRATION OF TEST EQUIPMENT

B′- FOR A LARGE MISCALIBRATION, FAILURE TO DETECT MISCALIBRATION FOR FIRST SETPOINT

C′- FOR A LARGE MISCALIBRATION, FAILURE TO DETECT MISCALIBRATION FOR SECOND SETPOINT

D′- FOR A LARGE MISCALIBRATION, FAILURE TO DETECT MISCALIBRATION FOR THIRD SETPOINT

Figure 7-4. Probability Tree Diagram of Hypothetical Calibration Tasks (small and large miscalibrations) will become suspicious and recheck his test setup. This is why the c path ends at S_2, a success endpoint in the event tree. However, given C, D = 1.0 means that if he does not become suspicious at the second setpoint, 100% of the time the technician will make the calibration change to the third setpoint without becoming suspicious. In routine tasks that are done frequently, this level of inattentiveness is not unusual. This pathway ends at F_1, one of the two possibilities for miscalibrating all three setpoints, the only failure of interest in the analysis.

When we move to the beta side of the tree, the recovery factors are better because the resultant changes to the calibration of the three setpoints will be substantial and should cue a competent technician. The reasoning for the beta side of the tree is the same as for the alpha side, except that the large calibration changes modify the estimates of conditional probabilities. Note that B' = 10^{-1}, which means b' = .9; i.e., 90% of the time the technician will be suspicious after calibrating the <u>first</u> setpoint. Given that he accepts this need for change, c' = .99 means that 99% of the time he will suspect a second setpoint requiring a large change. D' = 1.0 implies an unquestioning technician; if he accepts the first two changes, he will not question the third one. The values assigned to this pathway mean that such unquestioning behavior would occur only one time in a thousand, B' x C' X D' = $10^{-1} \times 10^{-2} \times 1.0 = 10^{-3}$.

Obviously, the above assumptions are not true in all cases. There is a possibility, for example, that the technician might recheck the test setup after becoming suspicious and conclude erroneously that the setup is correct, but we estimate this probability to be insignificant.

Analytic Development

The mathematical expressions for ZD and CD are fixed by definition. With ZD among n events, the equation for the probability of successful completion of all tasks is

$$\Pr[S|ZD] = ab\ldots n \tag{7-1}$$

where n is the probability of successful performance of the n^{th} activity, and the equation for the probability of incorrect performance of all tasks is

$$\Pr[F|ZD] = AB\ldots N \tag{7-2}$$

where N is the probability of failure on the n^{th} activity.

For the case in which the estimated failure (or success) probabilities for all n activities are equal,

$$\Pr[S|ZD] = a^{n} \tag{7-3}$$

where a is the probability of successful performance for the first of n events with equal estimated success rates, and

$$\Pr[F|ZD] = A^{n} \tag{7-4}$$

where A is the probability of error for the first of n events with equal estimated failure rates.

With CD among n events, the success equation is

$$\Pr[S|CD] = a \tag{7-5}$$

where a is the estimated success probability for the first action in a sequence of n completely dependent events, whether or not the basic success probabilities of each event are equal, and

$$\Pr[F|CD] = A \tag{7-6}$$

where A is the estimated failure probability for the first action in a sequence of n completely dependent events, whether or not the basic failure probabilities of each event are equal.

For the case in which it is not known which event will be performed first, but in which all events have an equal opportunity of being first, if the estimated success (or error) probabilities for the $\underline{n}$ events are not equal, the arithmetic mean of the $\underline{n}$ probabilities is used:

$$\Pr[S|CD] = \frac{a + b + \ldots + n}{\underline{n}} \tag{7-7}$$

where n is the estimated probability of successful performance of the $\underline{n}^{th}$ activity, and

$$\Pr[F|CD] = \frac{A + B + \ldots + N}{\underline{n}} \tag{7-8}$$

where N is the estimated probability of incorrect performance of the $\underline{n}^{th}$ activity.

In the probabiilty tree diagram (Figure 7-2, p 7-8), the above equations express the probability of arriving at the end of either the complete-success path or the complete-failure path. In the case of ZD, since there are no conditional probabilities involved, the success or failure estimates of each task are entered into the tree without modification. In the case of CD, only the two outer limbs of the tree are used, since the conditional probabilities of success and failure would both be 1.0, and the first branching inner limbs would have values of zero. For the intermediate levels of dependence, the conditional probabilities of success and failure for the inner branching limbs are obtained as described in the section "Dependence in Other Paths" (p 7-9).

The definition of the term basic human error probability (BHEP) is repeated here since this term is fundamental to calculating the conditional probabilites of all the limbs in a probability tree diagram. The BHEP is the probability of a human error on a task which is considered as an isolated entity, unaffected by any other task. Except for cases in which the first branching in a tree represents a carryover from another tree, the first failure limb in the tree is the BHEP of the first task considered. If ZD is assumed, the HEPs assigned to all the failure limbs in the tree represent BHEPs for the tasks in question. However, if the assumption of ZD does not hold, the BHEPs for succeeding failure limbs have to be modified to arrive at the appropriate conditional probabilities in accordance with the estimated level of dependence. In Figure 7-2, for example, the BHEPs for the limbs marked A and B|A might be equal, but, if CD is assumed, B|A becomes 1.0 rather than B. For the intermediate levels of dependence, the BHEPs are modified to obtain conditional probabilities, as explained later.

The complement of the BHEP is the BHSP, defined as 1.0 minus the BHEP. In cases in which dependence exists, the BHSPs also must be modified in accordance with the appropriate level of dependence.

Different Methods for Determining Intermediate Levels of Dependence

The conditional HEP (CHEP) of one task, given failure or success on some other task, is a function of the level of dependence between the two tasks. Our derivation of the intermediate levels of dependence results in the three intermediate points on the continuum of dependence as shown in Figure 7-1 (p 7-6).

These points were selected to divide the continuum of dependence among events that will lead to believable answers for the usual range of CHEPs in NPP operations. Several different approaches were considered, including the geometric mean approach used in WASH-1400. These approaches can be divided into two types -- one using nonlinear scaling and the other linear scaling.

In the former method, the CHEP is defined on a nonlinear scale that has the probability corresponding to ZD as a lower bound and the value 1.0 as an upper bound. One version of this method is a logarithmic scale. In this version the BHEP is expressed as BHEP = 10^x, where x is the exponent of 10 that yields the BHEP. For example, with a BHEP of 10^{-2}, then $x = -2$. For example, assume we have some basis to define the CHEPs as: 10^{x}, $10^{3/4x}$, $10^{1/2x}$, $10^{1/4x}$, and 10^{0} (or 1.0) for ZD, LD, MD, HD, and CD, respectively. Then, for a BHEP = .003 = $10^{-2.52}$, the CHEP for low dependence = $10^{-(3/4)(2.52)} = .013$.

There is a serious objection to this approach. When the BHEP is low, the method will generally yield unrealistically low joint HEPs (JHEPs). For example, if the BHEPs for two tasks to be done in succession are both .003 (our general BHEP of commission or of omission), the logarithmic method yields a JHEP for the tasks of .000039 ≈ .00004 for the low level of dependence. This probability is so small that one would often underestimate the influence of human errors.

The other major type of approach to dividing the dependence continuum, in which the CHEPs are almost independent of the BHEP, is to select values of dependence on a linear scale between the BHEP (ZD) and 1.0 (CD). In an earlier application the continuum of dependence was divided into four roughly equal parts. LD was about one-fourth the distance from ZD to CD, MD about halfway between ZD and CD, and HD three-fourths of this distance. In effect, this scheme yields estimates of .25 for LD, .50 for MD, and .75 for HD for the CHEPs of all events with BHEPs of .01 or less, given failure on the immediately preceding task. This method was discarded because the difference in JHEPs given ZD or LD conditions was much too large. For example, assume two tasks with BHEPs of 10^{-3} for each. Given ZD, the CHEP for a second task given failure on the first task would be the BHEP for that task; i.e., 10^{-3}, yielding a JHEP of 10^{-6}. Given LD, the CHEP would be approximately .25, and the JHEP would be 2.5×10^{-4}, a result very distant from the JHEP using ZD. It was therefore decided to consider a different set of values.

The division of the dependence continuum shown earlier in Figure 7-1 and the resultant values listed in Table 7-1 below provide more realistic reliability assessments and have been chosen as interim values for this issue of the handbook. Work continues in this difficult area of dependence, some of which is shown in the appendix to this chapter. The problem remains of developing a dependence model that meets certain statistical requirements and that provides realistic estimates of joint probabilities of failure based on human factors considerations. It is doubtful that this problem will be fully resolved.

The equations in the first column in Table 7-1 are used to calculate the conditional probabilities of success on Task "N," given success on the immediately preceding task, "N-1." The third column represents the conditional probabilities of failure on Task "N," given failure on the immediately preceding task, "N-1." The failure equations were selected so as to provide conditional probabilities of failure of 5%, 15%, and 50% of the distance between the BHEP (ZD) and 1.0 (CD), for, respectively, the low, moderate, and high levels of dependence. With BHEPs of .01 or smaller, the conditional failure probabilities will be approximately .05, .15, and .5. (The values for zero and complete dependence are, of course, fixed as the BHEP and 1.0, respectively.) The success equations were selected to provide an equivalent division of the dependence continuum between the BHSP and CD.

Application of Dependence Equations

To illustrate an application of the equations, refer to Figure 7-2 (p 7-8) and assume three tasks with BHEPs of, respectively, A = .001, B = .003, and C = .005, with LD between the first two tasks and HD between the second and the third tasks. To calculate the conditional probability B|A, Equation 7-15 from Table 7-1 is used:

Table 7-1. Equations for Conditional Probabilities of Success and Failure on Task "N," Given Success or Failure on Task "N-1," for Different Levels of Dependence

Success Equations	Equation No.	Failure Equations	Equation No.
$\Pr[S_{\text{"N"}} \mid S_{\text{"N-1"}} \mid ZD] = n$	(7-9)	$\Pr[F_{\text{"N"}} \mid F_{\text{"N-1"}} \mid ZD] = N$	(7-14)
$\Pr[S_{\text{"N"}} \mid S_{\text{"N-1"}} \mid LD] = \frac{1 + 19n}{20}$	(7-10)	$\Pr[F_{\text{"N"}} \mid F_{\text{"N-1"}} \mid LD] = \frac{1 + 19N}{20}$	(7-15)
$\Pr[S_{\text{"N"}} \mid S_{\text{"N-1"}} \mid MD] = \frac{1 + 6n}{7}$	(7-11)	$\Pr[F_{\text{"N"}} \mid F_{\text{"N-1"}} \mid MD] = \frac{1 + 6N}{7}$	(7-16)
$\Pr[S_{\text{"N"}} \mid S_{\text{"N-1"}} \mid HD] = \frac{1 + n}{2}$	(7-12)	$\Pr[F_{\text{"N"}} \mid F_{\text{"N-1"}} \mid HD] = \frac{1 + N}{2}$	(7-17)
$\Pr[S_{\text{"N"}} \mid S_{\text{"N-1"}} \mid CD] = 1.0$	(7-13)	$\Pr[F_{\text{"N"}} \mid F_{\text{"N-1"}} \mid CD] = 1.0$	(7-18)

$$\Pr[B|A|LD] = \frac{1 + 19(.003)}{20} = .05285 \approx .05$$

Since $B|A + b|A = 1.0$, $b|A \approx .95$.

To calculate the conditional probability, $C|B|A$, Equation 7-17 is used*:

$$\Pr[C|B|A|HD] = \frac{1 + .005}{2} = .5025 \approx .5$$

Since $C|B|A + c|B|A = 1.0$, $c|B|A = .4975 \approx .5$.

To calculate the conditional probability $b|a$, Equation 7-10 is used:

$$\Pr[b|a|LD] = \frac{1 + 19(.997)}{20} = .99715 \approx .997$$

Since $b|a + B|a = 1.0$, $B|a = .00285 \approx .003$.

To calculate the conditional probability $c|b|a$, Equation 7-12 is used:

$$\Pr[c|b|a|HD] = \frac{1 + .995}{2} = .9975$$

Since $c|b|a + C|b|a = 1.0$, $C|b|a = .0025$.

The remainder of the Task "C" limbs are calculated as follows:

for the expression	use the conditional probability	equals
c|b|A	c|b|a	.9975
c|B|a	c|B|A	.4975
C|B|a	C|B|A	.5025
C|b|A	C|b|a	.0025

To calculate the JHEP for the tasks that make up the complete-failure path in Figure 7-2 using the above conditional probabilities, take the product of the probabilities assigned to the three failure limbs:

$$\Pr[F] = (A \times B|A \times C|B|A) = .001 \times .05285 \times .5025 = 2.656 \times 10^{-5} \approx 3 \times 10^{-5}.$$

This value for the joint failure probability of these three actions, given the levels of dependence assumed, can be compared with the joint failure

---

*The calculation of the CHEP for a task is based on the influence of the immediately preceding task only.

probability that would be the product of the three BHEPs, given ZD, or 1.5 x 10^{-8}, an unbelievably low error probability for most human activities. Unless some reasonable approximation of dependence among human actions is included in estimating the impact of human errors on system reliability, one may grossly underestimate the effects of these errors.

To calculate the probability of the complete-success path in Figure 7-2 using the above conditional probabilities, take the product of the probabilities assigned to the three success limbs in that path:

$$\Pr[S] = (a \times b|a \times c|b|a) = .999 \times .99715 \times .9975 = .99366 \approx .994.$$

After the conditional probabilities have been assigned to all secondary limbs in the tree, the probability of any path through the tree is obtained by taking the product of the probabilities assigned to each limb in that path.

As indicated by the above examples, the level of dependence assigned between successive limbs in the tree can vary. The equations permit complete flexibility in this regard. Also, the BHEPs or BHSPs for the various tasks can differ. Although a factor of 5 difference was used in the above hypothetical example, usually the BHEPs are equal in a human reliability problem in which dependence effects must be estimated, or they differ by a factor of less than 3. However, the equations are valid even when there are gross differences in BHEPs. For example, assume HD between Tasks "A" and "B," and assume that the BHEP for the first task is .01 and for the second it is .001. The conditional probability B|A becomes

$$B|A = \frac{1 + .001}{2} = .5005 \approx .5.$$

Should the BHEPs for the two tasks be reversed, the equation for B|A results in an estimate of .505 $\approx$ .5. In both cases, the conditional probabilities of failure for Task "B" are essentially the same, but the joint probabilities of failure for the tasks differ by a factor of 10, depending on which task is performed first. That is, A x B|A = .01 x .5 = .005 for A = .01, but A x B|A = .001 x .5 = .0005 for A = .001.

From the equations in Table 7-1, it can be seen that as the BHEP gets smaller, the CHEPs for the succeeding tasks, given failure on the immediately preceding task, approach .05 for LD, .15 for MD, and .50 for HD. This was our purpose in selecting those particular equations. Therefore, it is suggested that these rounded CHEPs be used for BHEPs of .01 or lower. For BHEPs higher than .01, refer to Table 7-2 for the CHEPs or apply the equations from Table 7-1.

Table 7-2 also shows a desirable characteristic of the dependence model. Note that with lower BHEPs the ratio of CHEPs with LD to those with ZD grows larger (see the F columns). Psychologically, this has some advantages. As noted in Chapter 16, our lack of confidence in estimates of very low HEPs is reflected in the large uncertainty bounds assigned to such estimates. When one assumes ZD between two events with low BHEPs, this uncertainty is increased. For example, if the BHEP for each of two tasks is estimated as .003, our basic error of omission or commission, the joint probability of failure of both events, assuming ZD, becomes 9×10^{-6} (rounded to 10^{-5}), an unbelievably low figure for most human actions. If we substitute LD for ZD (usually a reasonable assumption), the resultant JHEP of 15×10^{-5} (rounded to 10^{-4}) is more realistic. In practice, of course, when the JHEPs are suspiciously low, we will question the estimates of the BHEPs, the level of dependence assessed, or both. In practical analyses, very low error probabilities indicate that the input data may be suspect. Although very low error probabilities may indeed reflect an excellent design, skepticism is usually warranted.

In contrast to the wide range of CHEPs, the range of values for CHSPs (given success on the immediately preceding task) will be restricted because

Table 7-2. Conditional Probabilities of Success or Failure for Task "N" for the Five Levels of Dependence, Given Failure on Task "N-1"

Task "N" Conditional Probabilities*

ZD**		LD		MD		HD		CD	
S	F	S	F	S	F	S	F	S	F
.75	.25	.71	.29	.64	.36	.37	.63	0	1.0
.9	.1	.85	.15	.77	.23	.45	.55	0	1.0
.95	.05	.9	.1	.81	.19	.47	.53	0	1.0
.99	.01	.94	.06	.85	.15	.49	.51	0	1.0
.995	.005	.95	.05	.85	.15	.50	.50	0	1.0
.999	.001	.95	.05	.86	.14	.50	.50	0	1.0
.9995	.0005	.95	.05	.86	.14	.50	.50	0	1.0
.9999	.0001	.95	.05	.86	.14	.50	.50	0	1.0
.99999	.00001	.95	.05	.86	.14	.50	.50	0	1.0

*All probabilities are rounded. Equations 7-14 through 7-18 were used to calculate the values in the F columns. The values in the S columns were obtained by subtraction.

**The conditional probabilities given ZD are the basic probabilities for Task "N."

Table 7-2

most tasks in NPPs have a BHSP of at least .99. As shown in Table 7-3, for BHSP of .99 or higher there is not much difference in the conditional probabilities for Task "N" regardless of the level of dependence assumed. However, note that Table 7-3 applies to the complete-success path in event trees, whereas one is usually interested in the complete-failure path and the listings in Table 7-2.

The effects of dependence on system reliability are not unidirectional. In some cases dependence between human activities will cause an increase in the overall probability of failure, while in others it will cause a decrease. The following observations on the effects of dependence apply to HEPs $< .5$ (almost always the case).

For the situations in which all tasks have to be performed without error for success (the complete-success path in an event tree), the JHEP decreases as the level of dependence increases. The magnitude of the decrease depends upon the size and relationship of the BHEPs. For example (using the values in Table 7-3), for three tasks with BHEPs of $A = .05$, $B = .01$, and $C = .001$, the overall probability of failure, $1 - abc$, is $1 - .95 \times .99 \times .999 \approx .06$ for LD $1 - .95 \times .995 \times .9995 \approx .055$ for HD. For BHEPs of $A = .001$, $B = .01$, and $C = .05$, the failure probabilities are $1 - .999 \times .99 \times .95 \approx .06$ for LD and $1 - .999 \times .995 \times .97 \approx .036$ for HD.

For the situation in which success is achieved if at least one task is performed without error (i.e., avoidance of the complete-failure path), the probability of failure increases as the level of dependence increases. For example (using the values in Table 7-2), for $A = .05$, $B = .01$, and $C = .001$, the probability of failure, ABC, increases from $.05 \times .06 \times .05 \approx 2 \times 10^{-4}$ for LD to $.05 \times .51 \times .50 \approx 10^{-2}$ for HD. For $A = .001$, $B = .01$, and $C = .05$, the probability of failure increases from $.001 \times .06 \times .1 = 6 \times 10^{-6}$ for LD to

Table 7-3. Conditional Probabilities of Success or Failure for Task "N" for the Five Levels of Dependence, Given Success on Task "N-1"

Task "N" Conditional Probabilities*

ZD**		LD		MD		HD		CD	
S	F	S	F	S	F	S	F	S	F
.75	.25	.76	.24	.79	.21	.87	.13	1.0	0
.9	.1	.9	.1	.91	.09	.95	.05	1.0	0
.95	.05	.95	.05	.94	.06	.97	.03	1.0	0
.99	.01	.99	.01	.991	.009	.995	.005	1.0	0
.995	.005	.995	.005	.996	.004	.997	.003	1.0	0
.999	.001	.999	.001	.999	.001	.9995	.0005	1.0	0
.9995	.0005	.9995	.0005	.9996	.0004	.9997	.0003	1.0	0
.9999	.0001	.9999	.0001	.99991	.00009	.99995	.00005	1.0	0
.99999	.00001	.99999	.00001	.999991	.000009	.999995	.000005	1.0	0

*All conditional probabilities are rounded. Equations 7-9 through 7-13 were used to calculate the values in the S columns. The values in the F columns were obtained by subtraction.

**The conditional probabilities given ZD are also the basic probabilities for Task "N."

$.001 \times .51 \times .53 \approx 3 \times 10^{-4}$ for HD. For more complex situations, the probability of failure may either increase or decrease with increasing dependence.

Application of the Dependence Model to a Parallel-Series System

To illustrate the application of the dependence model in a system context, two examples are offered following some general discussion. Both examples are based on the system shown in Figure 7-5. The three equipment components are manual valves, and the three human tasks are to restore these valves to their normal operating positions after maintenance. Any valve might fail to function properly on demand because of equipment defects or human failures. It is assumed that there is ZD between equipment and human failures, that there is ZD among equipment failures of the three valves, and that there is HD among the three human tasks, "A," "B," and "C," which are performed in that order.

For system success, the series leg (Valve #3) in Figure 7-5 and at least one of the parallel legs (Valve #1 or #2) must function properly. Therefore, there are three possible combinations of valve states leading to system success:

(1) All three valves succeed.

(2) Valve #1 succeeds, Valve #2 fails, and Valve #3 succeeds.

(3) Valve #1 fails, Valve #2 succeeds, and Valve #3 succeeds.

For any leg of the system, either a human failure, an equipment failure, or both will cause that leg to fail. Thus, from the logic of the system, seven possible success paths can be constructed as follows:

Success Path 1 = $E_1A \times e_2b \times e_3c$

Success Path 2 = $e_1A \times e_2b \times e_3c$

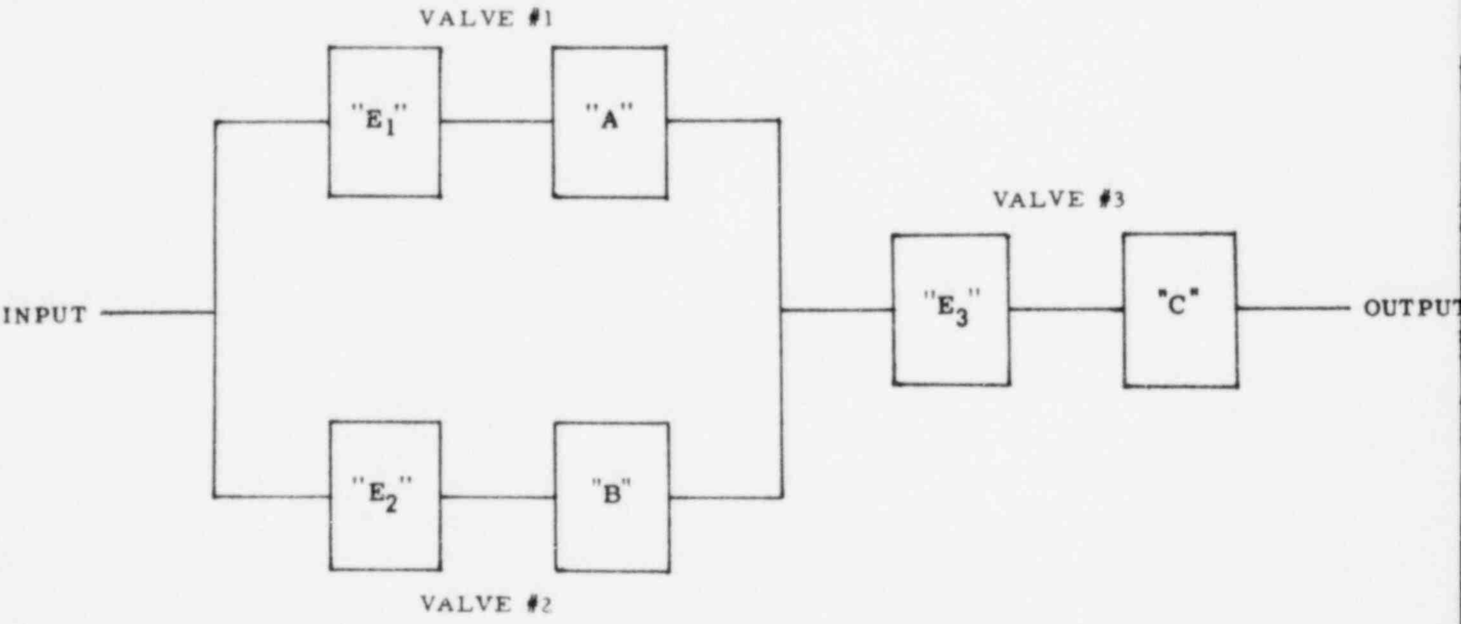


Figure 7-5. A Parallel-Series System with Potential Failures of Equipment (E_1, E_2, and E_3) or of Human Tasks (A, B, or C)

Success Path 3 = $E_1a \times e_2b \times e_3c$

Success Path 4 = $e_1a \times E_2B \times e_3c$

Success Path 5 = $e_1a \times e_2B \times e_3c$

Success Path 6 = $e_1a \times E_2b \times e_3c$

Success Path 7 = $e_1a \times e_2b \times e_3c$

For computational covenience, these seven paths can be reduced to three expressions that define the probability of system success as follows:

$$Pr[S] = Pr[S_1] + Pr[S_2] + Pr[S_3]$$

where $Pr[S_1]$ is the sum of success paths 3, 6, and 7 (equipment failures only or no failures); $Pr[S_2]$ is the sum of success paths 1 and 2 (human failures on the first parallel leg); and $Pr[S_3]$ is the sum of success paths 4 and 5 (human failures on the second parallel leg). Therefore,

$$Pr[S] = abc(E_1e_2e_3 + e_1E_2e_3 + e_1e_2e_3) + Ae_2be_3c(E_1 + e_1) + Be_1ae_3c(E_2 + e_2)$$

Since $(E_1 + e_1)$ and $(E_2 + e_2) = 1.0$, these terms can be dropped, leaving the following three expressions, with the terms rearranged to separate the human and equipment failure terms:

$$S_1 = abc(E_1e_2e_3 + e_1E_2e_3 + e_1e_2e_3)$$

$$S_2 = Abc \times e_2e_3$$

$$S_3 = aBc \times e_1e_3$$

In the examples below, we have calculated the probability of system failure under the assumption of ZD among human actions as well as under the assumption of HD among human actions, the original premise. Note that in one case the results are substantially the same whether or not dependence is considered, but that in the other case the difference is appreciable.

Example No. 1

For this example, we assume that the BHEPs for all three tasks are .01 and that the equipment failure probabilities are all .001. If ZD is assumed, the approximate failure equation is $Pr[F] \approx (E_1 + A)(E_2 + B) + (E_3 + C) = .011 \times .011 + .011 \approx .011$.

If HD is assumed for the human tasks, each success path must be calculated using values from Tables 7-2 and 7-3 for the human tasks:

$$\begin{aligned}
Pr[S_1] &= (a \times b|a \times c|b) \times [(E_1 \times e_2 e_3) + (E_2 \times e_1 e_3) + e_1 e_2 e_3] \\
&= (.99 \times .995 \times .995)\ [(.001 \times .999^2) + (.001 \times .999^2) + .999^3] \\
&= .97914
\end{aligned}$$

$$\begin{aligned}
Pr[S_2] &= A \times b|A \times c|b \times e_2 e_3 \\
&= .01 \times .49 \times .995 \times .999^2 = .00487
\end{aligned}$$

$$\begin{aligned}
Pr[S_3] &= a \times B|a \times c|B \times e_1 e_3 \\
&= .99 \times .005 \times .49 \times .999^2 = .00242
\end{aligned}$$

(It is conventional to postpone rounding of answers until the final answer is reached.)

The total failure probability for the system, including both human and equipment contributions, is

$$Pr[F_T] = 1 - \left\{ Pr[S_1] + Pr[S_2] + Pr[S_3] \right\} = 1 - .98643 \approx .014.$$

In this particular example, the practical effects of dependence were negligible (.014 vs .011).

<u>Example No. 2</u>

For this example, we will continue the assumption of .001 for the individual component failure probabilities and an HEP of .01 for Tasks "A" and "B," but will assume an HEP of .05 for Task "C." If ZD is assumed, the approximate failure equation is

$$Pr[F] \approx (E_1 + A)(E_2 + B) + (E_3 + C) = .011 \times .011 + .051 \approx .051.$$

If HD is assumed for the human tasks, the success paths are calculated as before, using the appropriate HD values for the BHEP = .05 from Tables 7-2 and 7-3 as follows:

$$\Pr[S_1] = (a \times b|a \times c|b) \times [(E_1 \times e_2e_3) + (E_2 \times e_1e_3) + e_1e_2e_3]$$

$$= (.99 \times .995 \times .97)\ [(.001 \times .999^2) + (.001 \times .999^2) + .999^3]$$

$$= .95454$$

$$\Pr[S_2] = A \times b|A \times c|b \times e_2e_3$$

$$= .01 \times .49 \times .97 \times .999^2 = .00474$$

$$\Pr[S_3] = a \times B|a \times c|B \times e_1e_3$$

$$= .99 \times .005 \times .47 \times .999^2 = .00232$$

The total failure probability for the system, including both human and equipment contributions, is

$$\Pr[F_T] = 1 - (.95454 + .00474 + .00232) = 1 - .96161 \approx .038.$$

In this example, failure to consider the effects of dependence yields a system failure estimate that is pessimistic by a factor of .051/.038 ≈ 1.3. In other cases, failure to use the appropriate level of dependence could produce overly optimistic results. There is no convenient method to determine beforehand whether dependence will have a substantial effect on system reliability. The effects of dependence should be calculated for each case, using the appropriate dependence equation, or, if data exist, a direct estimation of the level of dependence.

Sensitivity Analysis

Examples 1 and 2 illustrate sensitivity analyses in that several human performance measures were varied to determine their effects on overall system reliability. Example 1 showed that the difference in system outcome between the assumptions of ZD or HD was inconsequential. One can say, then, that the total system failure probability is insensitive to the level of dependence in the range of ZD to HD for the postulated HEPs. If the factor of 1.3 difference is considered significant, Example 2 illustrates that the system failure probability is sensitive to the levels of dependence assumed when the hypothesized HEP for Task "C" is increased from .01 to .05.

Finally, the two examples show how sensitive the system outcome was to an increase by a factor of 5 in the BHEP for Task "C." This increase in BHEP increased the total system failure probability from .014 to .038, roughly a factor of 2.7. In some analyses this increase would be considered negligible; in others, it could be important.

APPENDIX TO CHAPTER 7. AN ALTERNATIVE METHOD FOR ESTIMATING THE EFFECTS OF DEPENDENCE

By R. G. Easterling, Statistics, Computing, and Human Factors Division, Sandia National Laboratories

Introduction

As discussed in Chapter 7, there are two sources of dependence between the performances of two or more human tasks: (1) direct dependence, in which the performance of one task directly influences the performance of another task, and (2) common cause dependence -- the dependence of the performance of two or more tasks on common influences. This appendix sets forth the probability modeling required for expressing these two types of dependence.

Consider two tasks, "A" and "B," and let A and B denote the events: Failure on "A" and "B," respectively. Let C_i be the common condition present when "A" and "B" are performed. There is a collection $\left\{C_i\right\}$ of possible common conditions $i = 1,2,\ldots,\underline{n}$, and the event of interest is AB, failure on both tasks, <u>averaged</u> across the common conditions. For example, C_i might denote operator i. That is, the two tasks are to be performed by the same person. Different people have different skill levels and hence different probabilities of A and B. Over some period of time, say the plant's lifetime, "A" and "B" will be performed by different operators. Of interest, then, is the probability of A and B averaged over the population of operators.

As another example, suppose that because of training and the nature of the tasks it is reasonable to assume that all operators have the same basic skill levels. However, there may be a variety of variable influences (performance shaping factors) affecting performance so that a person's probability of A and B will vary, say from day to day. To obtain the overall probability of his failing on Tasks "A" and "B," this variability must be considered.

As yet another example, suppose C_i denotes plant i. Because of differences between training requirements, administrative procedures, plant designs, etc, the probability of A and B may vary from plant to plant. To obtain an "industry-wide" estimate of the probability of A and B, this variation must also be considered for. Not doing so can lead to underestimates.

Probability Model

Let $\Pr[A|C_i]$ denote the conditional probability of failure on "A," given condition C_i, and let $\Pr[B|A,C_i]$ denote the conditional probability of failure on "B" given condition C_i and failure on "A." It is through this latter probability, for each C_i, that the direct dependence of B on A is reflected (it is assumed that "B" follows "A"). Further, let $\Pr[C_i]$ denote the probability that condition C_i is present at the time "A" and "B" are performed. The C_is should be define such that they are mutually exclusive and exhaustive. That is, no two conditions can be simultaneously present, and all conditions must be included in the collection so that $\sum_{i=1}^{n} \Pr[C_i] = 1$. If it is possible for C_i and C_j to occur simultaneously, this overlap can be removed by defining three new conditions: (1) $C_1 = C_i$ alone present, (2) $C_2 = C_j$ alone present, and (3) C_3 = both C_i and C_j present. Exhaustion can be satisfied by defining C_n as "all other conditions." With this setup, the marginal (or average) probability of failure on both "A" and "B," denoted $\Pr[AB]$, is given by

$$\Pr[AB] = \sum_{i=1}^{n} \Pr[A|C_i] \Pr[B|A,C_i] \Pr[C_i]$$

This formula is a standard probability decomposition that results from the additive and multiplicative laws of probability. It reflects the dependence of A and B on the common cause, C_i. One special case worth noting is independence of both A and B on C_i. That is, suppose $\Pr[A|C_i]$ and $\Pr[B|A,C_i]$ do not depend on C_i. Then $Pr[A|C_i] = Pr[A]$ and $Pr[B|A,C_i] = Pr[B|A]$, and the above expression reduces to the more familiar

$$Pr[AB] = Pr[A]\ Pr[B|A],$$

which reflects the possible direct dependence of B on A. Only if it is reasonable to assume independence of A and B on C_i, that is, the absense of common causes, is it reasonable to approach estimation of Pr[AB] through this simpler expression.

Estimation of Pr[AB] through the above model may require considerable work by the analyst. The collection of conditions $\left\{C_i\right\}$ must be defined and then estimates of the $Pr[A|C_i]$, $Pr[B|A,C_i]$, and $Pr[C_i]$ must be obtained. It will be the exception, rather than the rule, that data-based estimates of these probabilities will be available. However, in many instances, a crude, subjective estimate of Pr[AB] may be all that is required, so that one may be able to use gross (or simplified) estimates of the required probabilities. For example, it may be possible to divide the set of conditions into three categories, say Good, Fair, and Poor, and estimate the corresponding probabilities or A and B. These conditions play the same role as PSFs. Through knowledge of the general variability of people and conditions and through analysis of the tasks of interest, one may arrive at reasonable and defendable estimates.

The probabilities likely to be the most difficult to estimate are the conditional probabilities, $Pr[B|A,C_i]$. The occurrence of A can be thought of as another PSF. Questions that must be addressed in estimating the effect of this factor in performing "B" include:

1. Is there complete feedback so that failure on "A" is known before doing "B?"
2. Is there incomplete feedback, as in the case in which one person may

know "A" was attempted, and assume it was done correctly even though it might not have been?

3. Is there physical linkage, as, for example, when "A" is to set dial A, and "B" is to line up dial B with dial A?

Note that dependence on a common cause is not considered here. That is, the concern is direct dependence of B on A, given a specific cause, or condition, C_i. The dependence due to common causes is reflected in the variability of $Pr[A|C_i]$ and $Pr[B|A,C_i]$ across the population of C_is, and this source of dependence is accounted for when the average of $Pr[A|C_i]\ Pr[B|A,C_i]$ across the C_is is calculated.

Converting answers to the above and similar questions to numerical values is the topic of this handbook. One useful approach is to begin with a "basic" set of probabilities $Pr[B|C_i]$, the conditional probabilities of failure on "B," given C_i but ignoring "A," and then adjust these probabilities depending on the psychological and physical linkages found between performance of the two tasks. One way to adjust the conditional probabilities is through consideration of the odds ratio. Let

$$L_i = \frac{Pr[B|A,C_i]/(1 - Pr[B|A,C_i])}{Pr[B|C_i]/(1 - Pr[B|C_i])}$$

The numerator of L_i is the odds ratio for the event B, conditional on A and C_i, while the denominator is the "basic" odds ratio. In considering how A shapes the performance of "B," one may plausibly think of the effect of A on the odds ratio. For example, given a fairly strong direct dependence of B on A, one might estimate that A increases the odds of B by a factor of 100, regardless of the condition C_i. That is, $L_i = 100$ for all i. Solving the above expression for $Pr[B|A,C_i]$ yields

$$Pr[B|A,C_i] = \frac{L_i Pr[B|C_i]/(1 - Pr[B|C_i])}{1 + L_i Pr[B|C_i]/(1 - Pr[B|C_i])}$$

At $\Pr[B|C_i] = .01$, for example, and $L_i = 100$,

$$\Pr[B|A,C_i] = \frac{100(.01/.99)}{1 + 100(.01/.99)} \approx .5$$

Lognormal Model

In some situations, rather than break down the common conditions into a few categories, it may be more appropriate to treat the C_is as an infinite set and the corresponding set of $\Pr[A|C_i]$s and $\Pr[B|A,C_i]$s as continuous random variables. For example, if C_i denotes operator i, rather than classifying operators as Good, Fair, or Poor in their performances of "A" and "B," it may be more appropriate to think of operator ability as a continuum. The mix of operator skill levels that might be called upon in performing "A" and "B" might then be represented by continuous probability distributions. For example, suppose $\log_e \Pr[A|C_i]$ and $\log_e \Pr[B|A,C_i]$ are assumed distributed (over the population of C_is) according to a bivariate normal distribution with means μ_A and $\mu_{B|A}$, variances σ_A^2 and $\sigma_{B|A}^2$ and correlation ρ_{AB}. Then

$$\Pr[AB] = \exp(\mu_A + \mu_{B|A} + \nu_A),$$

where,

$$\nu_A = (\sigma_A^2 + \sigma_{B|A}^2 + 2\rho\sigma_A\sigma_{B|A})/2.$$

(Note: This result assumes no truncation of $\Pr[A|C_i]$ and $\Pr[B|A,C_i]$ at 1.0.)

This model should be used only where there is strong justification. The discrete treatment of the C_is described previously provides an analysis for which the assumptions and their effects are much more transparent and hence more defendable.

CHAPTER 8. WALK-AROUND INSPECTIONS

One recovery factor used at nuclear power plants is the walk-around inspection (hereafter referred to as a "walk-around"). This is usually performed by an auxiliary operator or someone else designated by the shift supervisor. The walk-around is important because it allows plant personnel to detect some deficiency; e.g., an oil spill, a water leak, or a manual valve in the wrong position. Walk-arounds are especially important for detecting whether some ESF has been left unavailable after maintenance or testing.

While the walk-around offers one possibility of recovery from a human error, this recovery factor is not as effective as others because of the relatively passive nature of the inspection coupled with the operator's low expectancy of finding anything wrong. This chapter presents some performance models and estimates for the assessment of the recovery afforded by different applications of the walk-around, beginning with the basic walk-around and continuing with variations.

Basic Walk-Around

The basic walk-around consists of a scheduled inspection tour of a specified area in the plant. The operator* is merely told to report anything unusual or any deviant condition of equipment. If he is given a more explicit instruction (e.g., "Be sure to check the main isolation valve on the RWS the model for the basic walk-around is modified, as described later.

For the basic walk-around, we make the following assumptions:

(1) A walk-around is made once per shift.

(2) The operator performing the walk-around knows the plant well and can recognize deviant conditions if he notices them.

---

*In this chapter we will use the terms operator and inspector interchangeably.

(3) The operator covers the same area each time (although not necessarily in the same sequence), so that all <u>deviant items</u> have almost equal probabilities of being observed during each walk-around.

(4) No written procedure is used during the walk-around.

(5) No special oral instructions have been given to the operator to attend to some particular item of equipment.

(6) Any deviation is fairly obvious if the inspector knows what condition is normal. In this respect, the walk-around task is different from that of the inspector who looks for minor imperfections in product manufacturing, since the product inspector's judgments are more subjective. On the other hand, we are not addressing deviations as obvious as a large pool of water on the floor, which we assume will always be noticed.

The assumptions for this basic walk-around imply that the inspection is entirely visual and that the inspector is concerned primarily with things that are clearly deviant. They also imply that discovery of some undesirable situation (say, an oil slick on a floor) may interfere with his recognition of some other less obvious condition. If any of the above 6 assumptions is not valid, the HEPs in this section must be increased appropriately.

Given the above assumptions, we estimate a probability of .9 (.5 to .99) per exposure that an inspector will fail to notice a deviation (such as a manual valve in the wrong state) if there is no provision for indicating a disparity between the actual and correct states. This high probability of failure is due to a combination of PSFs.

The first of these is the inspector's set, or expectancy, to find important things as they should be. Expectancy is a very powerful PSF and a major cause of error in routine inspections when the probability of a deviant situation is low (e.g., equal to or less than 10^{-2}). Expectancy is based on past experience and the interaction between the inspector performing the walk-around and other inspectors and other personnel. This interaction includes the knowledge or feeling that maintenance and other personnel are responsible and competent and seldom make mistakes. In NPPs, the probability of there being a deviant item of the type that is looked for in a walk-around should be much less than 10^{-2} per shift. Therefore, the inspector performing the walk-around has a very strong expectancy of finding everything in order.

A second PSF derives from the decline of memory with time. If the valves present no indication of their normal status, the inspector must rely on memory. Even though the inspector understands the functions of all components in the plant, he tends to rely on visual memory when he makes his walk-around; i.e., if things look the way they looked the previous day, he will usually accept them as being correct (unless alerted, as by a deviation from an easily recognized pattern).

When a person has to inspect hundreds or thousands of items, his memory for individual items declines very rapidly. The curve of retention in Figure 8-1 is based on laboratory studies of visual recognition of words with the passage of time. The shape of the curve typifies other such studies and pertains to the inspection situation.

As shown, recognition accuracy one day after viewing a large number of items is only 25%. Our own estimate of 10% recognition accuracy on the first walk-around after the occurrence of a deviation is based on our assessment of a further degradation of 15 percentage points because of the strong expectancy to find all things normal in the plant. Assuming this degradation is constant and subtracting these 15 percentage points from the values in the curve, the probability of successful recognition on subsequent days will decline approximately as shown in Table 8-1. (A method of calculating the Pr[F] values is described in the appendix to this chapter.)

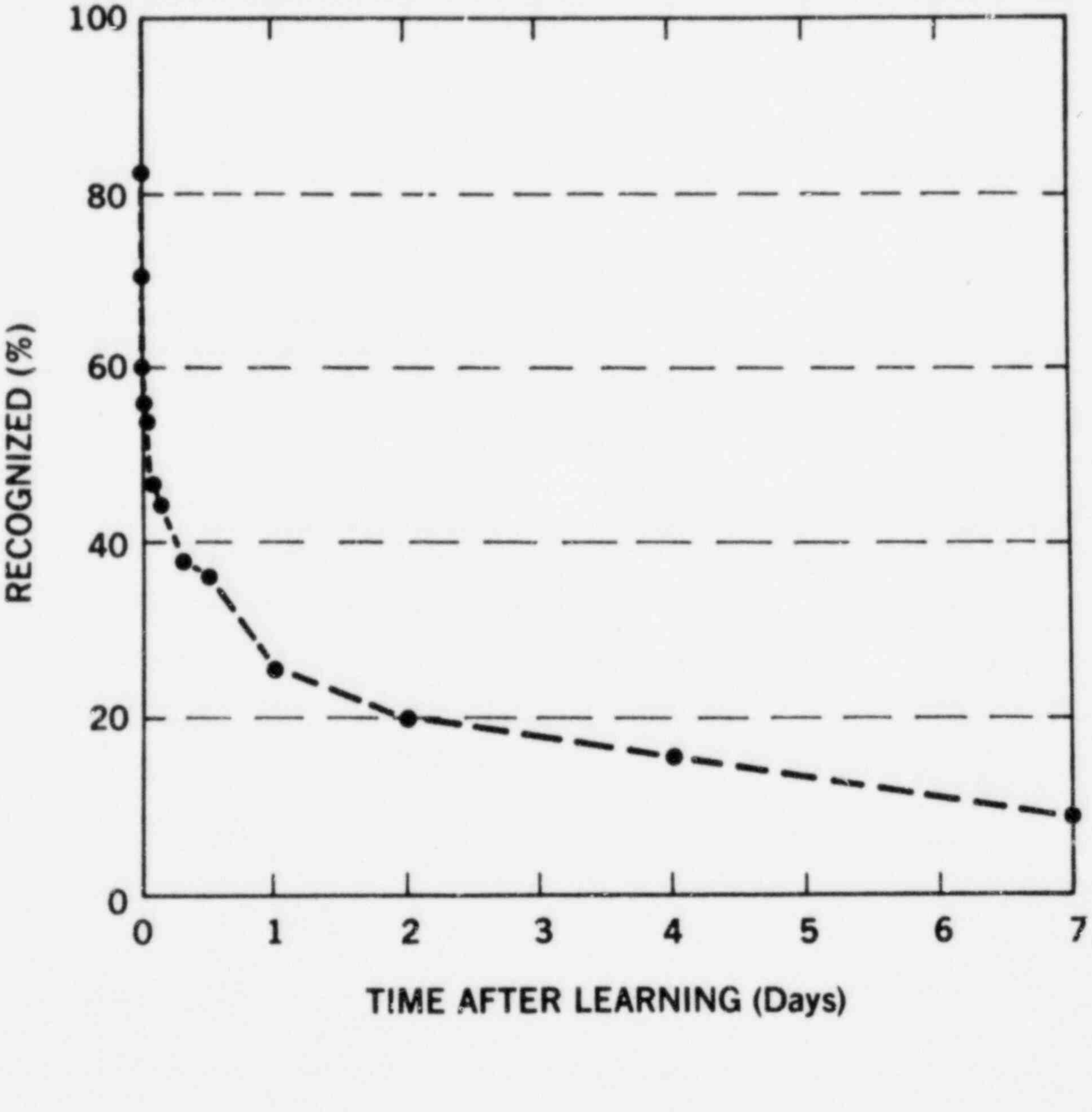


Figure 8-1. Curve of Retention as Determined by the Recognition Method (Woodworth, 1938, p 54)

Table 8-1. Declining Probability of Successful Recognition of a Deviant Item by One Person on Successive Days (Assuming One Walk-Around Per Day Per Person)

Day	Pr[S]	Pr[F]
1	.1	.9 (.5 to .99)
2	.05	.95 (.6 to .995)
3	.025	.975 (.7 to .999)
4	.01	.99 (.8 to .999)
5-30	.001	.999 (.9 to .999)
>30	0	1.0

NOTES: (1) The above estimates do not incorporate the special alerting effect of the detection of one deviant item on the detection of other deviant items during the walk-around, as discussed in the text.

(2) The Pr[F] on any given day is the probability of failure on that day given failure on all previous days.

(3) If a deviant item has been undetected for as long as 30 days, it is assumed that it will not be detected on subsequent walk-arounds unless some other indication alerts the operator to the original deviation.

The minimum Pr[S] of .001 on Day #5 (and 25 subsequent days) reflects a low-level recovery factor that has been observed in common experience in other activities. Let us assume that the inspector has seen a valve in the wrong state the day after it was placed that way, but did not note this deviation. His memory trace for the correct status declines as in Figure 8-1. At the same time, he develops memory traces for the valve in the wrong position, so that in time the new (incorrect) position becomes increasingly likely to be seen as correct. However, even under these conditions, people occasionally recognize an incorrect situation after having accepted it several times. We do not know why this happens; possibly the person is a little more alert than usual or he may just happen to be thinking of the logic underlying the valve position, but occasionally a recognition of the irregularity does occur after a succession of oversights.

Over a 30-day period, beginning with the first opportunity to detect a deviation after it has occurred, the total Pr[S] for an inspector who performs a walk-around once per day is:

$$\Pr[S_{\leq 30 \text{ days}} | \text{ 1 shift, 1 inspector}] \tag{8-1}$$

$$= 1 - \Pr[F_{\text{Day 1}}] \Pr[F_{\text{Day 2}}] \Pr[F_{\text{Day 3}}] \Pr[F_{\text{Day 4}}] \Pr[F_{\text{Day 5}}]^{26}$$

$$= 1 - (.9 \times .95 \times .975 \times .99 \times .999^{26})$$

$$= .1959 \approx .20$$

If different inspectors perform walk-arounds, the same Pr[S] applies to each. Thus, with the usual situation of one walk-around per shift, and given three shifts and the assumption of ZD between the shifts, the total probability of detection of some deviation over a 30-day period would be:

$$\Pr[S_{\leq 30 \text{ days}} | \text{3 shifts, 3 inspectors} | \text{ ZD between shifts}] \tag{8-2}$$

$$= 1 - \{1 - \Pr[S_{\leq 30 \text{ days}} | \text{1 shift, 1 inspector}]\}^3$$

$$= 1 - .8041^3 = .480 \approx .5$$

It can be argued that the assumption of ZD between the three shifts is overly optimistic. Some of our reviewers with operating experience suggest that the walk-around model should be modified for the evening and night shifts to take into account the likely dependence resulting from the attitude that "those guys on the day shift have probably found anything that I could find." This is a reasonable argument. However, since the walk-around HEPs are very large, the application of the dependence model will make very little difference. For example, if MD is assumed for the evening shift and HD for the night shift, Equation 8-2 is changed as follows, using Equations 7-16 and 7-17 (p 7-30):

$$\Pr[S_{\leq 30 \text{ days}} \,|3 \text{ shifts, } 3 \text{ inspectors}|\, \text{MD for 2nd shift, HD for 3rd shift}]$$

$$= 1 - .8041 \; \frac{1 + 6(.9)}{7} \times \frac{1 + 6(.95)}{7} \times \frac{1 + 6(.975)}{7} \times \frac{1 + 6(.99)}{7}$$

$$\times \; \frac{1 + 6(.999)}{7}^{26} \quad \frac{1 + .9}{2} \times \frac{1 + .95}{2} \times \frac{1 + .975}{2}$$

$$\times \frac{1 + .99}{2} \times \; \frac{1 + .999}{2}^{26}$$

$$= 1 - (.8041 \times .8302901 \times .8983408)$$

$$= 1 - .5997649 = .400 \approx .4$$

This example illustrates a case in which the assumption of ZD, although not valid from a psychological point of view, is accurate enough. In the remainder of this section we will assume ZD between shifts, since the calculations are simplified and no important accuracy is lost.

So, with the usual walk-around, there is about a 50-50 chance that any one deviation will be detected in a 30-day period. We assume that if a deviation remains undetected for 30 days after its occurrence, it will not be detected in subsequent walk-arounds. It is our judgment that if a deviant item remains undetected as long as 30 days, some factor is operating in the walk-arounds which prevents its being detected. Therefore,

$$\Pr[S_{>30 \text{ days}} \,|F_{30 \text{ days}}] = 0 \tag{8-3}$$

It is apparent from the above analysis that if each inspection could be made with a fresh inspector the probability of detecting a deviant item within a 30-day period would be very high. Assume, for example, the ideal but impractical situation in which each shift can draw inspectors from 30 operators, and in a 30-day period each operator performs only one walk-around, with a 30-day recovery period between walk-arounds for each operator. In such a situation, each operator has a first-day probability of detection of .1 each time he does the walk-around. Therefore, for any shift, the probability of successfully detecting a deviant item within 30 days becomes

$$\begin{aligned}&\Pr[S_{\leq 30\text{ days}}|1\text{ shift, }30\text{ inspectors}] \qquad (8\text{-}4)\\ &= 1 - \Pr[F_{\text{Day }1}]^{30}\\ &= 1 - .9^{30} \approx .96\end{aligned}$$

as compared with a Pr[S] of .2 for the same person performing the walk-around every day.

If we extend the above ideal to three shifts, requiring 90 different inspectors, the estimated probability of success is increased as follows:

$$\begin{aligned}&\Pr[S_{\leq 30\text{ days}}|3\text{ shifts, }90\text{ inspectors}] \qquad (8\text{-}5)\\ &= 1 - \{1 - \Pr[S_{\leq 30\text{ days}}|1\text{ shift, }30\text{ inspectors}]\}^{3}\\ &= 1 - .042^{3} \approx .9999\end{aligned}$$

(Note: If we assumed MD for the second shift and HD for the third shift, the .9999 estimate would change as follows:

$$\Pr[S] = 1 - \left[.9^{30} \times \left\{\frac{1 + 6(.9)}{7}\right\}^{30} \times \left\{\frac{1 + .9}{2}\right\}^{30}\right] = .99938 \approx .9994$$

Considering the inexactitude of the data, the difference between .9994 and .9999 may be disregarded.)

Of course, we do not assume that an NPP could afford to have 90 people available so that each of the 90 walk-arounds per month would be performed by a different person. Furthermore, it is not necessary from a risk-benefit standpoint. We estimate that if a person performs one walk-around per week, he will recover his recognition accuracy fully between inspections. Because of this recovery, we have a virtually independent inspection each week by that person. Our rationale for the full recovery in 1 week's time is based on common experience with the typical workweek. The performance of a chore once a week (with six intervening days of not performing the chore) essentially results in a fresh start for that chore each week. Thus, the estimated Pr[S] over a 30-day period of .9999 (from Equation 8-5) could be approached if each shift had seven people assigned to the walk-around so that any one person inspects only once every 7 days. This assignment would provide a 1-week recovery period for each operator.

The same type of analysis can be done for other intervals between walk-arounds if the relation between recovery and walk-around intervals is known. We assume that the recovery curve can be approximated by the cumulative normal distribution (with a mean of 4 and a standard deviation of 1). (For practical purposes, and as shown in the appendix to this chapter, the assumption of a straight line, a negatively accelerated curve, or a positively accelerated curve would not make any material difference in the calculations which follow.) Assuming the cumulative normal distribution and 100% recovery in 7 days, it is possible to construct the recovery curve shown in Figure 8-2.

Instead of seven inspectors assigned to a shift, consider a more realistic situation in which five people are available per shift and each person performs a walk-around every 5 days. From Figure 8-2, the recovery factor is .84. This recovery factor is applied to the decrement that would have occurred on Day #2 for the case of a single inspector making daily walk-arounds. In one month, any single inspector will make six walk-arounds, and his Pr[F] for any one deviant item, given previous nondetection, for each of them is shown below.

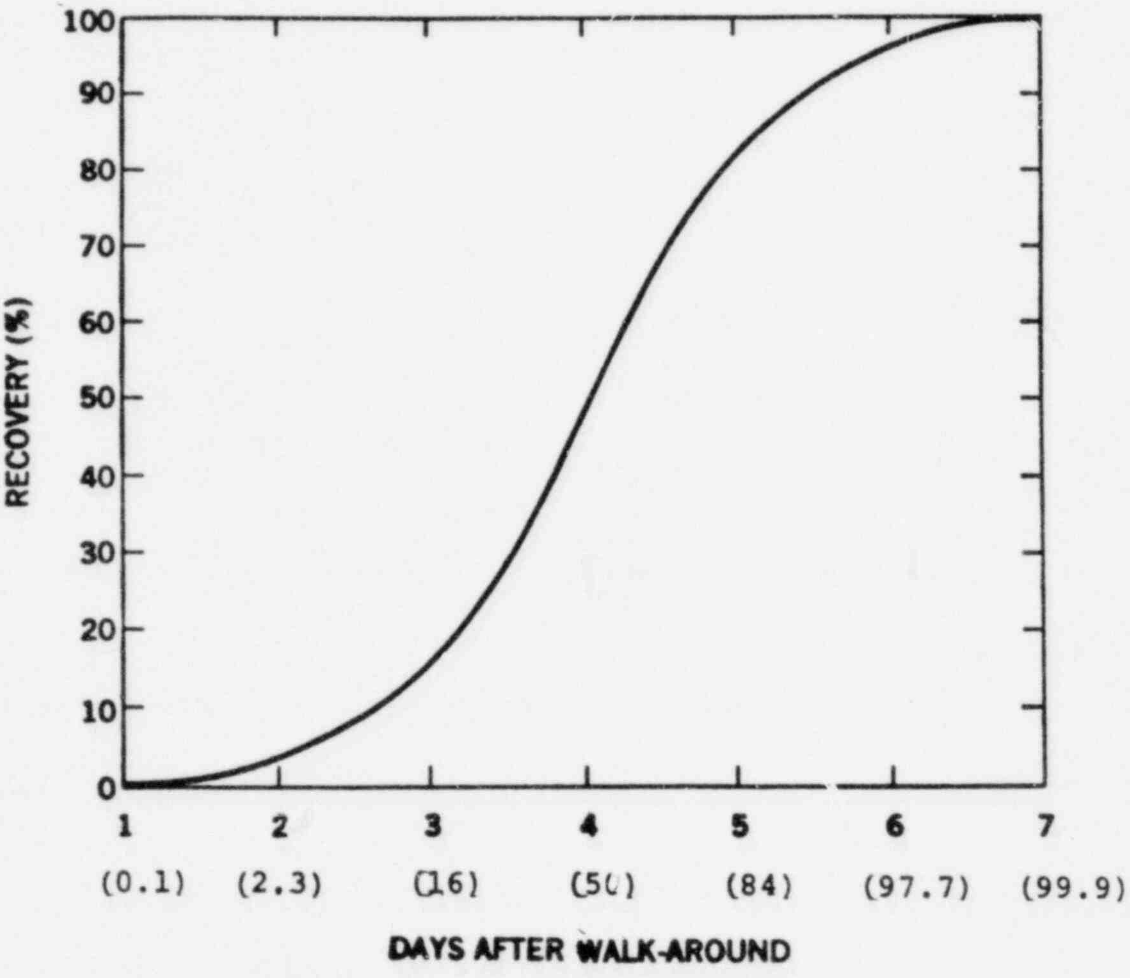


Figure 8-2. Percent Recovery by Days After Walk-Around

(The calculations are appended to this chapter.)

Walk-Around Number	Pr[F]
1	.9
2	.908
3	.915
4	.922
5	.928
6	.934

The Pr[S] over a 30-day period for one shift with five inspectors, each inspecting every 5 days, is

$$\Pr[S_{\leq 30} | 1 \text{ shift, } 5 \text{ inspectors}]$$
$$= 1 - [.9 \times .908 \times .915 \times .922 \times .928 \times .934]^5 = .92$$

For three shifts, the probability of detecting a deviation is:

$$\Pr[S_{\leq 30} | 3 \text{ shifts, } 15 \text{ inspectors}] = 1 - (1 - .92)^3 = .9995$$

Table 8-2 lists the probabilities of detecting a deviation over a 30-day period as a function of the number of inspectors that can be assigned to the same shift over the month's time. The first column in the table represents the number of inspectors assigned to a shift (with no provision for absence) and also represents the number of days from an inspector's first inspection to his next inspection. One walk-around per shift is assumed.

More Than One Walk-Around Per Shift

If different inspectors perform a walk-around in the same shift, their detection probabilities are combined. Thus, if two inspections per shift are performed by different inspectors (each performing one inspection per shift for 30 days), the probability of successful detection of a deviant condition within 30 days for one shift is calculated as follows, assuming ZD between inspections:

Table 8-2. Estimated Probability of Detecting a Deviant Condition in 30 Days or Less for Various Numbers of Days After Walk-Around*

Number of Inspectors Assigned to a Shift and Days After First Walk-Around	1 Shift	2 Shifts	3 Shifts
1 (daily walk-around)	.20	.35	.48
2	.35	.58	.72
3	.52	.77	.89
4	.77	.95	.998
5	.92	.99	.9995
6	.95	.9975	.9999
7 (weekly walk-around)	.96	.998	.9999

*Estimates of $\Pr[S_{\leq 30\text{ days}}]$ approaching .9999 should be carefully evaluated with respect to the assumptions beginning on p 8-1, especially #3.

$$\Pr[S_{\leq 30\text{ days}}|1\text{ shift, 2 inspectors, each making 1 inspection per shift}] \qquad (8\text{-}6)$$

$$= 1 - \Pr[F_{\text{Inspector 1}}|30\text{ days}]\ \Pr[F_{\text{Inspector 2}}|30\text{ days}]$$

$$= 1 - .8041^2 \approx .35$$

This value is the same as that for two shifts with one inspector each (Table 8-2). The Pr[F] of .80 is the complement of the Pr[S] for one inspector (.20).

If each of the three shifts uses the same procedure, the resultant Pr[S] for a 30-day period will be:

$$\Pr[S_{\leq 30\text{ days}}|3\text{ shifts, 6 inspectors, 2 per shift, each making 1 inspection per shift}] \qquad (8\text{-}7)$$

$$= 1 - \Pr[F_{\leq 30\text{ days}}|1\text{ shift, 2 inspectors, each making 1 inspection per shift}]^3$$

$$= 1 - .6466^3 \approx .73$$

This probability of .73 for two inspectors per shift can be compared with the probability of .5 for one inspector per shift for the detection of a deviation within 30 days (Equation 8-2). If the assumption of ZD is not valid and some higher level of dependence is more appropriate, the estimates of Pr[S] within 30 days will be lower. Assuming MD and using Equations 8-1 and 7-16 (p 7-30), the estimates of .35 for one shift and .73 for three shifts become, respectively, .33 and .70. The calculation of the .33 estimate is illustrated below:

$$\Pr[S_{\leq 30\text{ days}}|1\text{ shift, 2 inspectors, each making 1 inspection per shift}|\text{MD between inspectors}]$$

$$= 1 - \{[.9 \times \frac{1 + 6(.9)}{7}] \times [.95 \times \frac{1 + 6(.95)}{7}] \times [.975 \times \frac{1 + 6(.975)}{7}] \times [.99 \times \frac{1 + 6(.99)}{7}] \times \left[.999 \times \frac{1 + 6(.999)}{7}\right]^{26}\}$$

$$= 1 - .667634 \approx .33$$

One might presume that if an inspector performs a walk-around more than once per shift, there will be a net gain in the probability of his detecting a deviant condition. Even if we assume that the inspector's motivation will not suffer (not a realistic assumption), there will be no significant gain in the probability of detection over a 30-day period because the addition of 30 extra walk-arounds, with a probability of detection of .001 per walk-around, will change the overall .8 probability of failure (from Equation 8-1) by less than 3%. The calculation is:

$$\Pr[S_{\leq 30\ \text{days}} | \text{1 shift, 1 inspector making 2 inspections per shift}]$$

$$= 1 - (.9 \times .95 \times .975 \times .99 \times .999^{56}) \quad 1 - .78 = .22$$

$$\text{and } .02 \div .78 = .0256 < 3\%$$

Special Instructions to Inspector

Occasionally, the shift supervisor may ask the inspector to "be sure to check" some particular item of equipment to see if it is in the correct state. The term, one item of equipment, refers to a single item or a functionally related group of items that are completely dependent in that the inspector regards them as one unit and recalls them as one unit. (See the discussion of functionally related items beginning p 7-13.)

For the above particular item only, the Pr[S] is taken as .999 for the first walk-around inspection following the special instructions. Thus, Pr[F] is estimated as .001 (.0005 to .005). For subsequent walk-arounds, the data in Table 8-1 (p 8-5) hold since the special instructions no longer apply. Thus, the Pr[S] for the next day would be .1 for the same item. Although there might be some increase in attention over the next few days for an item singled out on one day, we have no estimate of this influence and, for conservatism, will ignore this effect. However, if the special instructions are repeated for the next walk-around, the .999 Pr[S] would again apply.

The following is our rationale for the .999 estimate. Although many studies of memory, or retention, have been conducted in laboratory settings, we have not found any published studies on retention of the above sort in the industrial situation. This is understandable since the typical industrial situation is arranged to function with minimal reliance on human memory.

Typical university laboratory studies test the ability of a person to remember relatively meaningless items such as numbers, syllables, or unrelated words (cf. Chapter X, "Retention and Forgetting," in McGeoch and Irion, 1952). The studies all yield substantially similar data: people can recall up to five items of the above kind for short intervals usually of no more than a few minutes.

However, meaningful material is retained for much longer intervals with relatively little decline over time. In the present case, we require the inspector to remember only one item, a meaningful one. This requirement is no more difficult than requiring a person to carry out an order. Given normally responsible personnel, deliberate failures to carry out orders are so rare that they may be disregarded as major sources of error.

Special instructions probably stimulate the inspector's interest in an otherwise routine walk-around. The only factor that might interfere with his retention of specialized instructions for one item of equipment would be some compelling distraction such as a major leak of some liquid or gas, a crack in a coolant pipe, a disturbing phone call from home, or some other unusual event that could "erase" the special instructions.

Considering the above factors and assuming routine plant conditions, we estimate that special instructions to attend to one particular item will be remembered and complied with, with an error of omission probability of about 10^{-3}. This estimate assumes that the instructions are clearly communicated and understood and that the inspector is fit for duty.

If the originator of the special instructions asks to be notified of the state of the equipment, the probability of an unrecovered error of omission is taken as zero. Our rationale is that even if the inspector forgets to check the item, he will be reminded of it when the originator asks for the information. Therefore, the Pr[S] is 1 minus the probability of a discrimination error (e.g., a valve is open, but the inspector makes a reversal error and "sees" the valve as closed). Discrimination errors of this type are discussed in Chapter 13.

The above discussion is for one item of equipment. Can we generalize from the laboratory data that suggest that people can retain five unrelated things in their memory stores? The differences in (1) required retention times, (2) the number of things that must be inspected in the walk-around, and (3) the number of competing stimuli, make such a generalization questionable. Our experience indicates that an individual cannot remember well more than two to three items. This applies to verbal communications as well as to items read from a list.

If no written list of the special items is made and the originator does not ask for feedback from the inspector, the basic Pr[S] of .999 for one item should be reduced as shown in Table 8-3 for additional items. Note that for the first three items the effect is minor, but increases rapidly thereafter. The values shown in Table 8-3 reflect the psychological truism that as the number of things to remember increases, the special instructions begin to equate to the meaningless instruction, "Look at everything!" Laboratory data show a somewhat faster decrease in ability to recall orally presented items,

Table 8-3. Probability of Recall of Special Instruction Items Given Orally as a Function of the Number of Items

Number of "Special Instructions" Item/Units	Pr[S] Recall of Any One Item	Pr[F] Recall of Any Given Item	Pr[S] Recall of All Items
1	.999	.001	.999
2	.997	.003	.994
3	.991	.009	.973
4	.973	.027	.896
5	.919	.081	.656

NOTE: Use uncertainty bounds of HEP ÷ 2 and HEP x 5.

but in these studies the items are meaningless (digits, nonsense words, etc) (cf. Woodworth, 1938, p 18). The values of Pr[F] in the table increase by a factor of 3 with each additional item to be remembered. This factor was selected because we judged it provided a reasonable modification of laboratory results for the added retention value of meaningful instructions.

Note that the table stops with five items. We think it unlikely that an inspector will be given more than five special items to check without his writing them down or being handed a list by his supervisor. Table 8-3 indicates that a written list should be used if more than two or three items are to be checked.

If the inspector is given a written list of special items to be checked, and if he checks off each item on the list as he finishes with it, his Pr[S] will be .999 for each item and $.999^{\underline{n}}$ for all $\underline{n}$ items. Obviously, if the list becomes long, say 10 items, some form of checklist should be considered, as described in the next heading. Also, if the inspector is asked to pay special attention to more than two or three items and is asked to report his results, it is presumed that he will write down the items to be checked rather than rely on his memory.

One final point on special instructions is mentioned here. When the inspector has been given special instructions to pay attention to some small set of items, his preoccupation with the special instructions may lower the probability of his noticing other possibly deviant items. Conversely, it can be argued that the special instructions have an arousal effect, and thus heighten the general level of vigilance and increase the probability of detection. Lacking data on the applicability of these two influences in NPP walk-arounds, we make the simplifying assumption that their effects cancel each other out.

Use of a Formal Checklist in the Walk-Around

If a formal checklist is used for a walk-around inspection, if it is used properly (as described below), and if the checklist indicates the correct status of each item to be checked, our estimate of the combined probability of looking at the item and recognizing an incorrect state is .99 -- HEP = .01 (.005 to .05).[*] This estimate assumes one walk-around per shift per person.

The proper way to use a checklist is to read each item in turn, inspect the item, then "check off" that item on the checklist, read the next item, and so on, continuously referring to the checklist as a guide to the sequence of inspections. Experience indicates that about half of the inspectors will use the checklist properly.[**] With increased familiarity, people take shortcuts: typically, an inspector will have the checklist with him and will inspect several items, check them all at once on the checklist, then check another "batch," and so on. This incorrect procedure increases the probability of errors of omission.

We assign a .9 probability of successful recognition of deviant items if the checklist is used improperly as descibed above.[***] Thus, given that the checklist will be used properly half the time and improperly half the time, the overall combined probability of looking at any item from the checklist and recognizing that it is in the incorrect state becomes

---

[*] For uncertainty bounds for the HEPs in the rest of this chapter, use HEP ÷ 2 for the lower bound and HEP x 5 for the upper bound.

[**] Some reviewers of the handbook have stated that our estimate of correct use of a checklist 50% of the time may be optimistic.

[***] We have also observed personnel going through an entire walk-around, returning to their normal work location, getting out the checklist, and then checking off all the items. If this practice is followed at a plant, no credit is allowed for the "use" of a checklist; i.e., use the values from Table 8-1 (p 8-5).

$$\Pr[S_{\text{for any deviant item}} | 1/2 \text{ proper use and } 1/2 \text{ improper use of checklist}] = .5 \times (.99 + .9) = .945 \approx .95 \qquad (8\text{-}8)$$

The Pr[S] for all n deviant items is calculated by taking the arithmetic mean of the estimated performances of the two types of inspectors:

$$\Pr[S_{\text{all } n \text{ deviant items}} | 1/2 \text{ proper use and } 1/2 \text{ improper use of checklist}] = .5 \times (.99^n + .9^n) \qquad (8\text{-}9)$$

Thus, if there were as many as five deviant items in a long checklist (a most unlikely situation), the probability that all such deviant items would be recognized is calculated as

$$\Pr[S] = .5 \times (.99^5 + .9^5) \approx .77$$

or, about three-fourths of the time all five deviant items would be found.

If a person performs more than one walk-around per shift, it is likely that his use of the checklist will become casual or perfunctory. We estimate a .9 probability of success per deviant item per inspection under these circumstances.

If the checklist does not indicate the correct state of each item, it will still serve as a reminder to look at each item. In this case, errors of discrimination as well as errors of memory must be considered.

No extra credit is allowed for a requirement that the inspector initial each entry rather than merely check it off. A mark is a mark. The continual use of initials, and even signing at the end of a checklist, becomes perfunctory. We do recommend signing each checklist, however, since this information can be useful for other purposes.

If two people are assigned to the walk-around, one as a reader and the other as a checker (as discussed in Chapter 15, "Recovery Factors and Administrative Control"), there should be some increase in the probability of recognition of deviant items of equipment. The gain is attributed to the increased probability that the checklist will be used correctly rather than to the incorporation of "an extra pair of eyes," because we assume high dependence between the team members. The probability of recognition of deviant items for two people is therefore taken to be .99.

If two inspectors are assumed without a checklist, we assign no credit for the second person. We estimate that this situation is equivalent to a case of complete dependence, and the values in Table 8-1 (p 8-5) apply.

APPENDIX TO CHAPTER 8. CALCULATIONS FOR WALK-AROUND Pr[F]s AS A FUNCTION OF PERIOD BETWEEN SUCCESSIVE WALK-AROUNDS

Figure 8-2 (p 8-10) lists the hypothesized percentage of recovery (and the corresponding Pr[F]s) as a function of days after a walk-around inspection. The recovery applies to the loss in detection probability that occurs if walk-arounds are conducted daily as shown in Table 8-1, p 8-5. For example, on Day #1 the Pr[F] is .9, and on Day #2 it is .95 -- an increase of 5 percentage points. These values must be adjusted to allow for the fact that walk-arounds may not be performed on a daily basis. The percent degradation in performance is affected by the percent recovery that occurs as a function of the days between the inspections performed by any one operator. If 5 days elapse after the first inspection, the recovery factor of .84 (from Figure 8-2) is applied to the 5 percentage points (as described above, the expected daily performance degradation from the first to the second inspections), reducing the degradation to D = .05 - (.84 x .05) = .008 and yielding a Pr[F] of .9 + .008 = .908 for the second trial.

The amount of degradation on any day immediately following a walk-around is a function of the Pr[F] for that day. As shown in Table 8-1, when the initial Pr[F] is .90 the degradation is 5 percentage points, when it is .95 the degradation is 2.5 percentage points, and when it is .999 there is no further degradation until the limiting Pr[F] = 1.0 after 30 days. We assume a linear relationship between D (the degradation that occurs on Day $\underline{n} + 1$) and $\Pr[F_{\underline{n}}]$ (the probability of failure on Day $\underline{n}$). The equation below expresses this function:

$$D = .5 \times (1 - \Pr[F_{\underline{n}}]) \qquad (8\text{-}10)$$

where D = the increase in Pr[F] on Day$_{\underline{n}+1}$. In calculating the Pr[F] for a trial subsequent to any given trial, the complement of the recovery factor from Figure 8-2 is multiplied by the degradation calculated for that subsequent trial, and the result is added to the Pr[F] for the first trial. Thus, for any trial $\underline{n}$, the Pr[F] for the subsequent trial, $\underline{n}$ + 1, by the same inspector, is

$$\Pr[F_{\underline{n}+1}] = \Pr[F_{\underline{n}}] + D \times (1 - r) \qquad (8\text{-}11)$$

where r is the recovery factor based on the number of days between successive trials (from Figure 8-2). As an example, if we assume an interval of 5 days between inspections, r = .84. The Pr[F] for Day #1 is specified as .9, and the subsequent Pr[F]s are calculated as follows:

Day #	D	r	D x (1 - 4)	Pr[F$_{\underline{n}}$]
1	.05	.84	.008	.9
2	.046	.84	.007	.908
3	.0425	.84	.0068	.915
4	.039	.84	.0063	.922
5	.036	.84	.0057	.928
6				.934

From Equation 8-10, D for Day #1 is .5 x (1 - .9) = .05, and from Equation 8-11, Pr[F] for Day #2 is .9 + .05 x (.16) = .908. For Day #2, D = .5 x (1 - .908) = .046, and Pr[F] for Day #3 is .908 + .046 x (.16) = .915. The calculations are continued through Day #6, which will complete a 30-day cycle with five operators. The Pr[S] for the 30-day period is 1 - $(.9 \times .908 \times .915 \times .922 \times .928 \times .934)^5$ = .92.

For the case of one operator conducting an inspection every day, the calculated probabilities will be slightly different from the ones listed in

Table 8-1. Table 8-1 indicated a Pr[F] of .999 by Day #5, whereas Equation 8-11 would yield the following Pr[F]s:

Day #	Pr[F]
1	.9
2	.95
3	.975
4	.988
5	.994
6	.997
7	.998
8	.999

The difference in the Pr[S]s for a 30-day period is negligible, .1959 using the values in Table 8-1, and .2039 using the calculated values above, both of which round to .20.

In the earlier example involving five inspectors, the number 5 was a convenient factor of 30. The following example illustrates the calculations with other numbers. Assume 1 shift, four inspectors, each inspecting at 4-day intervals. Each will make seven inspections, and two will make an additional inspection to complete the 30-day period. Therefore, we calculate Pr[F]s for 8 days, using an r of .5 (from Figure 8-2).

Day #	Pr[F]	D	r	D(1 - r)
1	.9	.05	.5	.025
2	.925	.0375	.5	.01875
3	.944	.028	.5	.014
4	.958	.021	.5	.0105
5	.968	.016	.5	.008

Day #	Pr[F]	D	r	D(1 - r)
6	.976	.012	.5	.006
7	.982	.009	.5	.005
8	.987			

The Pr[S] for 30 days is calculated as follows:

$$\Pr[S_{30\ \text{days}}] = 1 - [(.9 \times .925 \times .944 \times .958 \times .968 \times .976 \times .982)^4 \times .987^2] = .77$$

The estimated probabilities of detecting a deviant condition in 30 days or less for different intervals and numbers of shifts are listed in Table 8-2. The Pr[F]s underlying the figures in Table 8-2 are listed below.

Pr[F] per Trial as a Function of Interval Between Trials

Number of Days Between Trials, and Recovery Factors (from Figure 8-2)

Trial #	2(2.3%)	3(16%)	4(50%)	5(84%)	6(97.7%)
1	.9	.9	.9	.9	.9
2	.949	.942	.925	.908	.901
3	.974	.966	.944	.915	.902
4	.987	.980	.958	.922	.903*
5	.993	.989	.968	.928*	.904
6	.996	.993	.976	.934	
7	.998	.996	.982*		
8	.999	.998	.987		
9		.999			
$\Pr[S_{30}]$ =	.35	.52	.77	.92	.95

*This number of trials completes a 30-day cycle.

The above figures are all based on the assumption of a recovery curve that follows a cumulative normal distribution. It was also mentioned that the assumption of some other distribution would not have much practical effect on the calculations. The Pr[F]s per trial were recalculated using the assumption of a straight-line recovery, yielding the values below.

Pr[F] per Trial as a Function of Interval Between Trials, Assuming Straight-Line Recovery Functions

Number of Days Between Trials, and Recovery Factors

Trial #	2(17%)	3(33%)	4(50%)	5(67%)	6(83%)
1	.9	.9	.9	.9	.9
2	.942	.933	.925	.917	.909
3	.966	.956	.944	.930	.916
4	.980	.970	.958	.942	.923
5	.988	.980	.968	.951	.930
6	.993	.987	.976	.959	
7	.996	.991	.982		
8	.998	.994	.987		
9	.999	.996			
10		.997			
$Pr[S_{30}]$ =	.44	.60	.77	.87	.93

The estimates listed in Table 8-2 were then recalculated, and are listed in Table A8-1.

Clearly, the inspection model is not very sensitive to differences in the assumptions of the shape of the recovery function. Either assumption indicates that acceptable levels of probability of detection can be achieved by judicious trade-offs between the number of shifts and the number of inspectors assigned to each shift (to permit variation in the number of days between inspections for each operator).

Table A8-1. Estimated Probability of Detection of a Deviant Condition in 30 Days or Less Assuming a Straight-Line Recovery

Number of Inspectors Assigned to a Shift and Days After First Inspection	1 Shift	2 Shifts	3 Shifts
1	.19	.34	.47
2	.44	.69	.82
3	.60	.84	.94
4	.77	.95	.988
5	.87	.98	.998
6	.93	.995	.9997
7	.96	.998	.9999

CHAPTER 9. DISPLAYS IN CONTROL ROOM

This chapter presents some of the background and assumptions for estimating errors of omission and commission in reading various types of displays in NPP control rooms. Specific performance models for annunciated and unannunciated displays are presented in Chapters 10 and 11, respectively. Although these chapters specifically address displays in control rooms, many of the statements and estimates apply equally well to displays in other areas. The user of the handbook must assess the relevant PSFs for the application and make extrapolations accordingly.

No attempt has been made to differentiate among the different design versions of any given type of display with regard to their influence on human error. For example, the typical fixed-scale, moving-pointer analog meter may be in vertical, horizontal, curved, or other form. While these meters may differ in their susceptibility to reading errors, the data are so sparse that it is not possible to substantiate such differentiation.

In general, NPP displays are not optimal for reducing human error. Our performance models are based on current, typical designs in LWRs and may be conservative (by about a factor of 2 to 10) when applied to displays that conform to accepted human factors design practices.

Some Basic Assumptions

There are hundreds of displays in a typical NPP control room. In addition to unannunciated displays such as indicator lights, analog displays, digital readouts, and computer printouts, there are annunciated indicators that are typically transilluminated legend lights equipped with auditory alarms and blinking signals to gain the operator's attention. When the auditory and blinking signals have been turned off, the annunciated indicator operates like any other unannunciated legend light. Indicator lights are generally of two types. One is a capped, covered lamp with a label above or below the lamp. The other is a legend light (a transilluminated label). Analog displays are generally charts, meters, or recorders. Meters take several forms: vertical, horizontal, circular, etc, generally with fixed scales and moving pointers. Graphs are also used, but usually in books or procedures rather than on a control room panel. Digital readouts generally display three or four digits. Computer printouts may be displayed on video screens or may be printed out by typewriter.

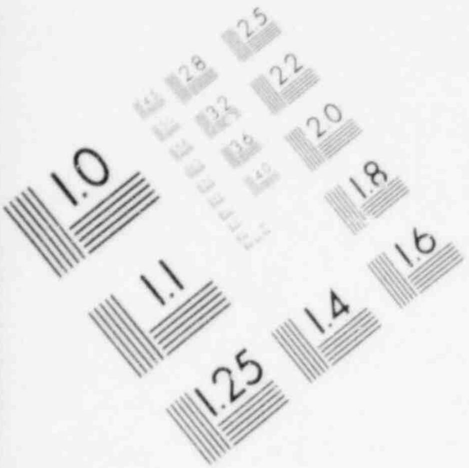
In most control rooms, the annunciators are grouped in arrays above the other control/display panels and are considerably above eye level. The status lamps, digital readouts, and analog displays are mounted on a variety of panels -- some vertical, some horizontal, some slanted. The layout of the control/display panels in NPP control rooms often increases the complexity of the operator's job and reduces the reliability of his performance.

When a new shift begins, it is assumed that the oncoming operator will conduct an initial survey or audit of all the control boards and will then monitor them throughout the shift to note any deviant displays or deviant manual controls. Since his primary interest is in maintaining the supply of power to the grid, he will pay most attention to displays directly related to this function. Thus, although he will not deliberately ignore displays related to other functions (e.g., safety-related displays), he will allot less scanning and monitoring time to them. (The low incidence of failure indications in NPPs naturally lowers his expectancy of finding any deviant safety-related displays.)

At some plants, the control room operator must manually log about 15 to 20 specific parameters every 2 hours. Some of these parameters relate directly to system safety; e.g., pressurizer level. At other plants, manual logging is not required and reliance is placed on computer printouts. Obviously, the second practice reduces the likelihood that the operator will detect some deviant parameter before it causes an annunciator to alarm.

In this handbook, the primary interest is in safety-related equipment and systems. The estimates of scanning efficiency in Chapters 10 and 11 pertain to safety-related displays that are not used in the moment-by-moment running of the plant. We assume that scanning and monitoring efficiency is less for solely safety-related displays than for displays related to keeping the plant online. Even though most safety-related functions are annunciated, it is important for control room personnel to scan related unannunciated displays to anticipate potential trouble and take preventive action as early as possible. If a safety-related function is moving toward some value at which automatic equipment normally takes over, the alert operator will be readier to initiate manual action in case of automatic equipment failure than if his first warning of trouble is an annunciator.

In the case of annunciated displays, the annunciated indication will very likely be detected when it comes on, and some action will be initiated within a few minutes at most. Similarly, an unannunciated display which is directly related to an annunciated indication will have a very high probability of being checked. If an annunciator sounds, the operator will usually check the annunciated indicator and then check the related unannunciated displays. We judge that the model for detecting annunciated displays holds for other directly related displays. By "directly related," we mean that the labels on the annunciated legend lamp, in effect, direct the operator's attention to a specific display or displays. For example, if an annunciated legend lamp states that the pressure for some function is low, the operator would be expected to refer to the pressure display for that function. The use of the



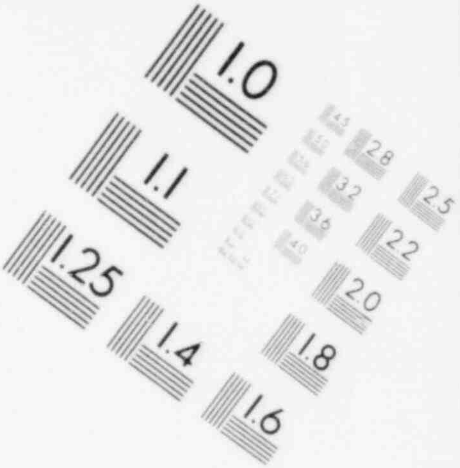
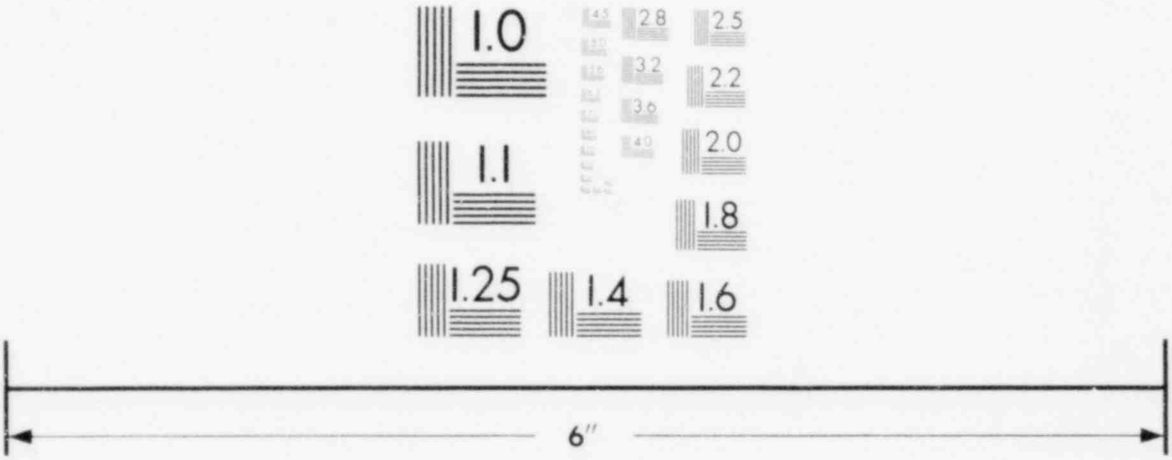
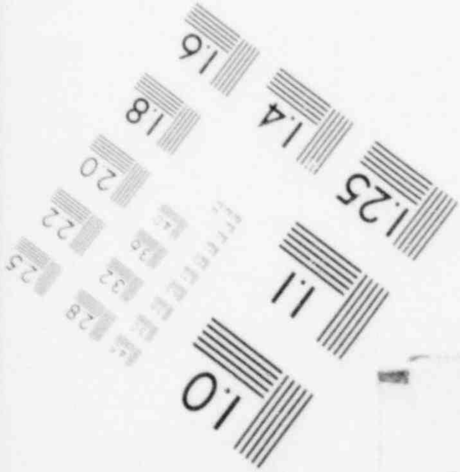
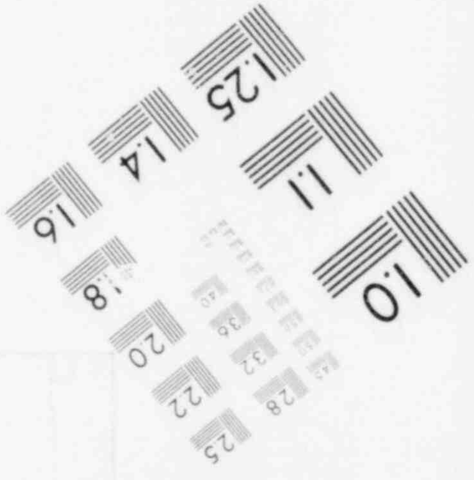


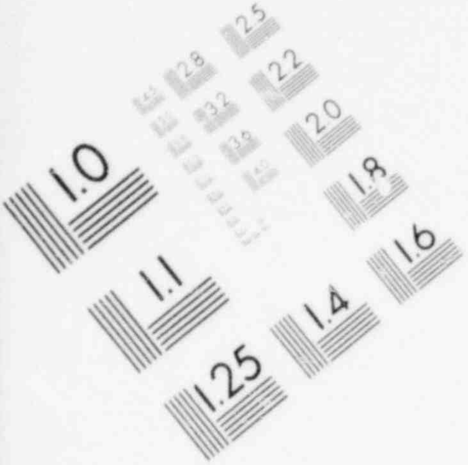
IMAGE EVALUATION TEST TARGET (MT-3)



MICROCOPY RESOLUTION TEST CHART







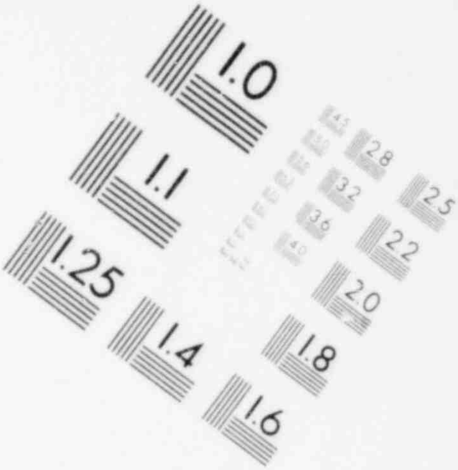
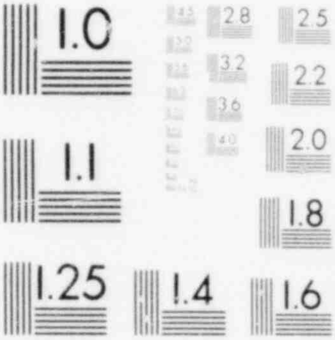
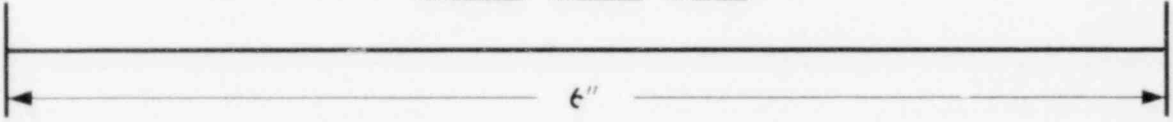
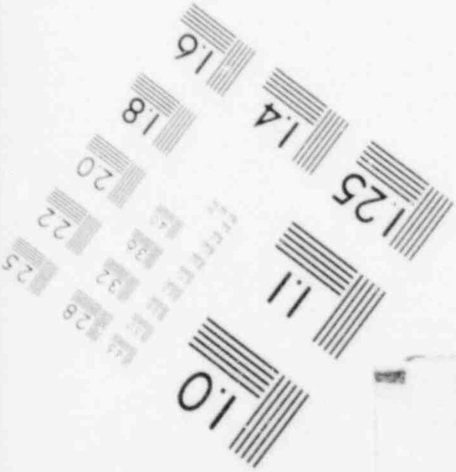


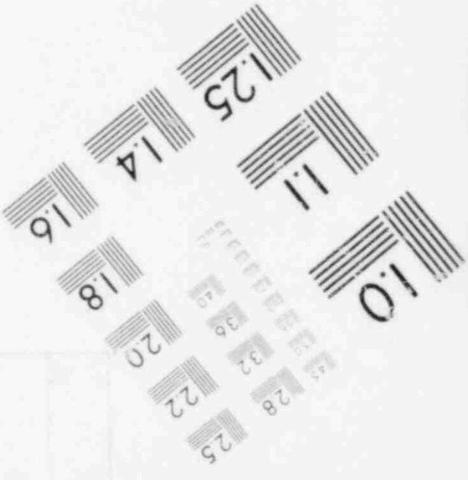
IMAGE EVALUATION
TEST TARGET (MT-3)





MICROCOPY RESOLUTION TEST CHART





annunciator model rather than the assumption of CD allows for some possibility of distraction or other interruption of the operator's task.

For routine detection of unannunciated deviant displays, reliance must be placed on the frequency and reliability of the scanning patterns used by control room personnel. The models for unannunciated displays presented in Chapter 11 are based on the assumption of hourly scanning as representative of the scanning habits of control room operators for certain types of displays. For other types, an assumption of one check per shift (usually during the initial audit) is made. For still other displays, such as those related to nonchanging functions, the assumption of even one check per shift may be optimistic. These assumptions are based on interviews with control room personnel.

One Deviant "Steady-State" Display - One Operator

If we consider a single deviant safety-related display in a steady-state condition* that was not detected in the previous shift, the oncoming shift operator will have some probability of detecting that display during his initial scan of the control boards. This probability varies according to the type of display. We hypothesize that, for any given initial probability of detection, the probability of detection over the entire remainder of the shift is best represented by an exponential curve. We further assume that this relationship holds for the case in which an operator regards a particular functional group of displays as a single unit.

Since we are assuming trained and experienced operators, it is reasonable to expect that they will know what is associated with what. For example, if

---

* A "steady-state" condition of a display means that meters or digital readouts are not rapidly changing status, lamps are not blinking, and the auditory signals for annunciators are canceled. These displays present no special alerting cue to the operator. The steady-state condition of a display is different from the steady-state operating mode of the plant, which refers to the normal power generating condition.

an operator notes an increase in containment temperature, he will usually check the containment pressure. These two indicators are judged to be completely dependent. There are many other cases in which more than one display may be treated as a single unit when estimating the probability of detection. We are not able to formulate specific rules for what constitute a single unit since plants differ. The principle is that closely associated indications in a system may be regarded as single units. The models for displays pertain to individual indicators and to any group of indicators that constitute a single psychological unit.

Under highly stressful conditions, especially when immediate action is required, some operators may not respond to a whole unit (e.g., both temperature and pressure displays), but may fixate on one element of the unit. Several incidents have occurred in which an operator did not cross-check directly related instruments but concentrated on one instrument. In some cases, the one instrument displayed erroneous information or insufficient information to enable the operator to interpret the situation correctly.

It is general policy in NPPs that outgoing control room personnel brief the incoming shifts. Interviews and observation reveal much variability in the thoroughness of this intershift consultation. This is in part a function of what problems the outgoing shift has experienced, and in part it is a function of personalities. Ideally, the incoming control room operator would perform his initial audit in the company of the outgoing control room operator. Then the incoming operator would continue this audit for some time beyond the departure of the outgoing operator. This is the period when the probability of detection of any deviant condition is highest. For calculational convenience and with negligible loss of accuracy, we assume that the initial audit takes place at the very beginning of the shift and that its duration is zero time.

Based on interviews with operators, we can identify three kinds of steady-state displays:

(1) those which characteristically are not checked during a shift (e.g., a pair of status lamps that indicates whether some blocking valve is open or closed),

(2) those which are sufficiently important that they would usually be checked once in the shift (e.g., the level in the Refueling Water Storage Tank), and

(3) those on which important information may change frequently, requiring that they be observed several times per shift (e.g., containment temperature).

For the first kind of steady-state display, we assume no periodic scans. For the second kind, we assume that one scan occurs during the initial audit. For the third kind on which important information may change frequently, we hypothesize an initial scan, followed by hourly periods of scanning if nothing unusual occurs. We assume that scanning effectiveness for such displays decreases for the rest of the shift because of the cumulative effects of fatigue, boredom, and expectancy, with an end spurt in effectiveness coincident with shift turnover to the oncoming operator.

The shape of the distribution of scanning accuracy is conjectural because the lack of data, the variation among operator scanning habits, and the variety of possible situations in a plant. As shown in Figure 9-1, we have selected the exponential curve as most closely representing the decline of scanning effectiveness after the initial scan and through the seventh hourly scan. For calculational convenience, we have incorporated the end spurt effect in the estimate of detection effectiveness during the initial audit in the succeeding shift.

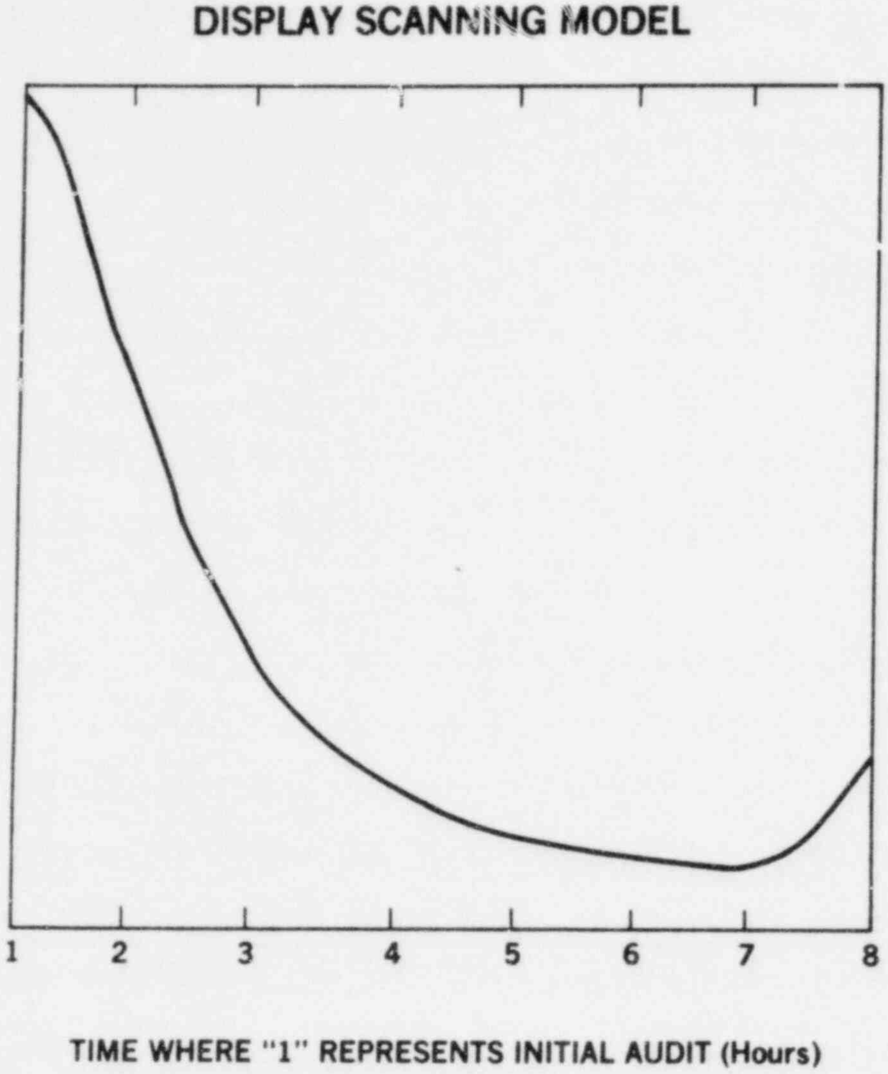


Figure 9-1. Hypothetical Curve Representing Probability for Detection Effectiveness At and Following Initial Audit (8 scans are shown at hourly intervals, beginning with initial audit and ending with last hourly scan in shift)

During the initial audit, the probability of the operator's detecting is highest because he is fresh and unaffected by the recent history of the displays. If he misses a deviant indication in the initial audit, he will be less likely to notice it on the next scan because he has already accepted it as normal. With each successive failure to detect, the probability of detection on the subsequent scan decreases as the operator grows more accustomed to the deviant indication.

In cases involving an indication that becomes deviant after the initial audit, there are two opposing psychological influences. We expect that the highest probability of detection will occur on the first scan after that change, with this probability of detection declining exponentially on subsequent scans. However, there is also an expectancy effect that will reduce the probability of detection of a deviation that occurs after the initial audit. Consider, for example, a deviation that occurs midway in the shift; i.e., just before the fifth scan. Having seen "good" indications in the previous four scans, the operator expects to see a good indication on subsequent scans as well. This is analogous to the situation of an inspector on a production line where quality is very good; he seldom experiences a bad unit, and, expecting them all to be good, frequently misses the occasional bad item.

We do not know how to quantify the effects of either influence on behavior. We will assume they cancel each other, and will consider the exponential curve in Figure 9-1 to represent the instantaneous probability of detection of any previously undetected deviant display.

When the probability of detection of a particular type of deviant display is very low, a further simplification is made without significant loss of accuracy in a human reliability analysis. For displays with a Pr[S] of about .01, we simply ignore the exponential curve and assume a constant probability of successful detection per scan. Our reasoning is that when the initial probabilities are so low, the use of the exponential curve constitutes questionable exactitude.

One Deviant Display - Two Operators

If two operators are assigned to the control panels, and if both operators scan for deviant indications, the probability of detection should be greater than with only one operator. This section presents our estimates of the added effectiveness that can be expected from the use of more than one operator.

Several factors must be considered:

(1) number of operators assigned to the control panels of a single reactor,

(2) division of the task; for example, one operator scans the reactor control board and the other scans the rest of the panels,

(3) degree of interaction in a multiple control room among operators who are assigned to different reactors,

(4) level of dependence among the operators assigned to a single reactor, and

(5) percentage of time the operators are available to scan the panels.

Number of Operators

It is possible that more than two operators might be assigned to monitor the control panels of a reactor, but we doubt that this would add materially to the probability of detecting deviant items. Usually, the third operator would be assigned to other duties, either formally or informally, and only two operators would be effectively scanning the panels. We assume a maximum of two effective operators. If more than two control room operators are assigned to one reactor per shift, we recommend that only two be assumed in a human reliability analysis.

Assignment of Panel Responsibilities

Seminara et al (1976) state in their review that when two operators were assigned to a control room for a given reactor, they were not assigned to different panels. That is, they did not divide the work. We assume the same practice in this handbook.

For the case in which a second operator in a control room is assigned a particular function or functions, we assume that his reliability for those functions equals that of the regular control room operator for the rest of the control room tasks. For example, following the accident at TMI, some plants have assigned a dedicated operator whose primary function is to maintain sufficient water in the steam generators. In general, his basic reliability should be the same as that of the regular operator. However, the estimated HEPs for the relatively few tasks performed by a dedicated operator can be materially reduced by the elimination of the interpretations and decisions characteristically required of the regular operator. (This case is discussed further beginning p 21-14.)

Interaction Among Operators Assigned to Different Reactors

At some plants, a single room will house the control panels for two or three reactors, each with its own assigned control room operators. Under normal operating conditions, we assume no interaction among operators of adjacent reactors even though the panels are in the same room. However, under unusual conditions it is likely that an operator from one reactor will assist the operator(s) of another reactor. This is discussed in Chapter 17, "Stress."

Level of Dependence Between Two Operators Assigned to Same Reactor Panels

If both operators are active monitors in the sense that both are equally responsible for the panels, we judge that MD would best reflect the interaction between them when both are performing their duties in the control room. If this condition is not met, the use of MD will result in an underestimate of the joint probability of their failing to detect a deviant display.

Another possibility would be to have one active operator and one relatively passive operator assigned to the control room. In such a case, a high level of dependence would be a more appropriate estimate of their interaction. If the second operator functions as a "rover" (that is, checking on things outside the control room), he is not available for control room activities during his absences.

Percentage of Time Operators are Available to Scan the Panels

In deriving the estimated probabilities of detection of deviant displays, we did not assume that an operator would be constantly scanning the control boards. This would be unrealistic since operators vary in the frequency and thoroughness with which they scan the control boards. We have assumed no scans, one scan, or hourly scans for different types of steady-state displays during one shift.

We judge that when two operators are present and both have been instructed to scan the same control boards, each will show less scanning effectiveness than if he were the sole operator. This was the rationale for the assumption of MD. Also, it is likely that because of the imposition of extra tasks when two operators are assigned, each will have less time for scanning than if he were the sole operator. The question is: For what percentage of the shift will two operators be available for scanning and for what percentage will only one operator be available? Instances during which only one operator is available for scanning include any time during which only one operator is present in the control room as well as those times when one operator is engaged in paperwork or other activity that precludes effective scanning. In the absence of any direct measurement, we suggest use of a 50-50 split. This means that half of the time MD is assumed, and half of the time only one operator is present.

Multiple Deviant Displays - One Operator

If more than one display becomes deviant during a shift, there is obviously a greater probability that at least one of them will be detected, since there are more chances to see a deviant display.* In practice it is unlikely that there will be more than five unannunciated deviant displays in an NPP. We judge that with six or more deviant displays there will be some type of auditory annunciation. We have developed a general model for annunciated displays, described in Chapter 10. The rest of the present chapter deals with unannunciated displays only.

The extent to which each additional deviant display facilitates detection is a function of the detectability of the individual displays. For example, a deviant meter attracts much more attention than does a deviant status lamp. Thus, there is a higher probability of detecting at least one of several deviant meters than of detecting at least one of several deviant status lamps. Because of the differences in the cumulative facilitative effects of signals with different basic probabilities of detection, it was necessary to develop a detection model that allows for these differences. Application of the concept of dependence satisfied this requirement, and is in accordance with what is known of human performance in NPPs. The application of dependence to

---

*It is assumed that the deviant indication will remain deviant until something is done about it.

detection is premised on the statistical relationship between some set of events and some influencing factor -- a common cause influence (p 7-2). The most important common cause influence affecting detection of deviant indications is a facet of human performance over which we have very little control -- the scanning habits of the operator. For a variety of reasons, operators will pay more attention to some displays than to others, and will be more responsive to some displays than to others. The practical effect of differences in the PSFs of displays is identical to the effects of direct dependence among events, and may be analyzed similarly. For the detection model, the application of HD among events yields realistic estimates of HEPs as a function of the number of deviant displays. Table 9-1 lists probabilities of detecting at least one deviant display when there are from one to five such deviant displays. The entries are obtained by use of the following equation, which is derived from Equation 7-17 (p 7-30). Equal BHEPs for the deviant displays are assumed.

$$\Pr[S_{\text{one or more deviant displays}}|\text{equal BHEPs}] \tag{9-1}$$

$$= 1 - \Pr[F_{\text{one}}]\left[\frac{1 + \Pr[F_{\text{one}}]}{2}\right]^{\underline{n} - 1}, \quad \underline{n} \leq 5$$

where n is the number of deviant displays.

This equation can be modified for the case in which one wishes to estimate the probability of detection of at least one display from a set of deviant displays for which all the BHEPs are not equal. Such application should be made only within the same class of displays; e.g., to meters or to chart recorders or to some combination of status lamps and legend lights. Thus, for the case of four deviant meters, two with limit marks and two without limit marks, two equations would have to be used since one would not know the order of displays to be scanned, and the BHEPs for these two types of

Table 9-1. Estimated Probabilities of Detecting At Least One of *n* Deviant Displays

	BHSP for Number of Deviant Displays				
BHEP	1	2	3	4	5
.99	.01	.015	.02	.025	.03
.95	.05	.07	.10	.12	.14
.9	.1	.15	.19	.23	.27
.8	.2	.28	.35	.42	.48
.7	.3	.41	.49	.57	.63
.6	.4	.52	.61	.69	.75
.5	.5	.63	.72	.79	.84
.4	.6	.72	.80	.86	.90
.3	.7	.81	.87	.92	.95
.2	.8	.88	.93	.96	.97
.1	.9	.95	.97	.98	.99
.05	.95	.97	.99	.993	.996
.01	.99	.995	.997	.999	.999

NOTE: For HEPs $< .5$, the lower uncertainty bound is calculated as HEP $\div$ 5 and the upper bound is HEP x 2. For HEPs $\geq .5$, the lower uncertainty bound is $1 - 2(1 - \text{HEP})$ and the upper bound is $1 - 0.2(1 - \text{HEP})$.

meters are different. One equation would assume that a meter with limit marks is scanned first and the other equation would assume that a meter without limit marks is scanned first. The arithmetic mean of the probabilities of success obtained from the two equations would be used as the estimate of the probability of detecting at least one of the four meters. (The BHEPs for meters and other displays are listed in Chapter 11.)

Under normal conditions, when an operator detects a deviant display he will also check functionally related displays. The probability of detecting related deviant displays follows the model for annunciators -- the first detected deviant display is treated as the equivalent of the first of several annunciators, and the related deviant displays as the equivalent of the rest of a set of annunciators. In applying this rule, it is the operator's perception of what is related that must be considered. The TMI accident teaches us that operators do not necessarily perceive the functional relationships that would enable them to cope with unusual events. Differences in training and experience of the operators will materially affect this perception. No detailed guidelines can be given for applying or not applying the above rule; this is a judgment to be made by the analyst.

In addition, there will be an arousal effect that will heighten sensitivity to other, unrelated deviant indications. For these displays, we assume that the arousal will raise the detection effectiveness to the level at the time of the initial audit. The arousal effect will last until the operator is satisfied that everything is back to normal. For simplicity, we assume instantaneous arousal and dissipation.

Multiple Deviant Displays - Two Operators

Equation 9-1, for calculating the probability that an operator will detect at least one of up to five deviant displays of the same class (see p 9-13), should be modified for the case in which there are two operators assigned to the panels of a reactor. If we assume that the two operators are both scanning for half of the shift, their joint probability of failure to detect at least one of up to five deviant displays would be:

$$\begin{aligned} \Pr[F_{\text{both operators}}] &= .5 \times \Pr[F_{\text{both operators}} | \text{MD}] \\ &\quad + .5 \times \Pr[F_{\text{one operator}}],\ \underline{n} \leq 5 \end{aligned} \qquad (9\text{-}2)$$

where $\underline{n}$ is the number of deviant displays of the same class, both operators are available 50% of the time, and only one operator is available the other 50% of the time.

CHAPTER 10. ANNUNCIATED DISPLAYS

This chapter presents a model describing the probabilities of errors of omission for scanning and reading of annunciated displays, also referred to as annunciators. The model applies to four situations in the control room: (1) steady-state operating mode (i.e., the normal power generating condition in the plant), (2) maintenance or calibration operations involving control room personnel, (3) anticipated transients, and (4) loss-of-coolant accidents (LOCAs).

Types of Annunciated Displays

In NPPs, most annunciated displays are incorporated in several panels of legend lights above the vertical control boards, above eye level. There are from 15 to 66 lights (or tiles) per panel, and there may be 400 to 750 such lights per reactor. When any annunciated function deviates from a specified condition, an automatic signal initiates an auditory alarm (horn, bell, or buzzer) and causes one or more of the legend lights to blink. Separate buttons cancel the auditory and blinking signals. When the blinking signal is canceled, the indicator remains illuminated in a steady-on condition until the trouble is cleared. Generally, 20 or more lights are on at any one time due to various conditions that do not require immediate action.

At most plants, when the trouble is cleared auditory and visual "clear" signals occur that differ from those signaling the onset of the problem. The "clear" signals may then be canceled with the appropriate button.

There also are auditory alarms for automatic printout equipment. Finally, there are auditory high-radiation alarms, fire alarms, security alarms, and so on. This chapter deals only with annunciated legend lights and printout equipment.

Following are the major types of possible error:

1. Annunciated legend lights
   - a. Omission Errors: failure to initiate some kind of corrective action within a few minutes (this may include an incorrect decision to take no action)
   - b. Errors in scanning of unannunciated conditions of annunciated legend lights: during any subsequent scan, failure to recover the initial error of failing to respond to the steady-on legend light
   - c. Reading Errors: either the wrong light is read or an error is made in reading the correct light
   - d. Judgment Errors: the indication is read correctly but the operator makes the wrong decision as to the required action
2. Annunciated printout equipment
   - a. Omission Errors: same as 1.a above, but this error may include a failure to read the printout
   - b. Errors in Scanning: same as 1.b above
   - c. Reading Errors: the message is incorrectly read
   - d. Judgment Errors: same as 1.d above

The performance model and estimated HEPs for responding to annunciated indicators are speculative and may be modified when objective data so indicate. In the interim, we have taken a conservative position in assessing the reliability of operators in responding to annunciated indicators. The basic problem is that one can expect a wide variety of responses because of the very large number of annunciated indicators and the fact that the indicators do not always provide the precise information the operator needs for making timely and correct decisions. Among experienced operators, errors of judgment unrelated to situational causes (such as poor display of information and lack of onsite practice) are so infrequent that their probabilities can be included in the error bounds assigned to the HEPs for the related tasks. (Errors of judgment are further discussed in Chapter 19, p 19-12.)

In future plant designs, if sound ergonomics practices are followed (cf Seminara, Eckert, et al, 1979), large improvements in operator reliability can be expected because many of the factors contributing to operator error will be eliminated. Regular onsite practice of simulated emergency conditions also improves operator reliability. The model and HEPs in this chapter reflect the inadequate human engineering of present plants and the inadequate practice that operators receive in dealing with emergency conditions.

Major Performance Shaping Factors

The most important PSFs are: (1) the number of signals per unit time that the operator must process, (2) the number of relatively unimportant indicators, (3) the number of false alarms, (4) the placement and design of annunciator systems and indicators, and (5) the stress levels (ranging from boredom to the emotions associated with a major accident).

Table 10-1 lists some human engineering deficiencies related to annunciator warning systems that affect the above PSFs. The table is based on the human reliability analysis in WASH-1400, on subsequent NPP studies in the U.S. and Europe, and on the EPRI Review (Seminara, et al, 1976).

Errors in Reading Annunciated Legend Lights

When an annunciator light comes on, the operator normally cancels the sound, looks to see which light is blinking, cancels the blinking, and reads the message on the legend light. Two major reading errors are possible. First, when the operator looks away from the blinking light to find the cancel button for the blinking, he may look at the wrong steady-on light when he

Table 10-1. Human Engineering Deficiencies
in Annunciator Warning Systems (Page 1 of 2)

1. The large number of annunciator lights per reactor (400 to 750) coupled with the large number of alarms even under normal operating conditions (estimated by EPRI interviewees as from 2 to 30 times per hour, depending on the situation) call for complex response patterns. The discrimination requirements can be excessive.

2. The number of false alarms per shift (estimated by EPRI interviewees as ranging from 15 to 50%, with as many as 75 false alarms in some cases) leads to the expected reaction to a "cry wolf" situation.

3. The normal background (20 or more) of annunciator lights in the steady-on state reduces the signal-to-noise ratio for an operator searching for meaningful displays.

4. Critical warnings are interspersed with noncritical warnings. This reduces the arousal effects of the former, and makes them more difficult to identify.

5. The lack of location aids as well as the large number of illuminated annunciator lights contribute to the fact that some operators have experienced difficulty in finding some safety-related annunciators during simulated accidents.

6. In some control rooms, large viewing distances are combined with small lettering on indicators. This requires that the operators move around quite a bit to read all the indicators. Some operators try to identify an alarm by its position on the board rather than by approaching it to read its label.

7. The intensity of auditory alarms is, in some cases, loud enough to evoke a startle response. This motivates operators to silence the alarm immediately, sometimes without even looking up at the annunciator boards. The intensity is so compelling as to interfere with the task at hand and cause forgetting. In some cases, coins are used to lock HORN SILENCE buttons in the cancel position. In some plants, the operator may have to leave a panel where he is performing some critical task to silence an auditory alarm. Operators have been observed to lose track momentarily of what was going on.

8. Some annunciated indicators tell the operator that either a high or a low setpoint has been exceeded, but not which one. In some cases, an annunciated indicator means that any one of four possible conditions exists.

9. There is a lack of uniformity across plants in the relative locations of the HORN SILENCE button, the ACKNOWLEDGE button (which turns off the blinking of the lights and leaves them in a steady-on condition), the LAMP RESET button (which turns off the lights when the trouble has been cleared), and the TEST button (which causes all lamps to light and blink while the button is depressed). (Different plants may have different designations for these buttons.)

Table 10-1 (p 2 of 2)

10. Acknowledgment of an annunciated legend light causes its annunciator to go to a steady-on state and to blend in with all the others in the steady-on state. This can result in a loss of information.

11. The difference between the Alert and Clear blink rates is not always immediately apparent to the operators.

12. The alarm audio frequency may be too high for some operators to hear.

13. Simulator experience shows that, in a major accident, annunciators come on in such bewildering numbers that it is not possible to read, much less absorb, the meaning of all the annunciators. In some cases, they are deliberately ignored (Kemeny, 1979).

14. According to some simulator instructors, the primary motivation of even skilled operators (undergoing recertification exercises) during simulated abnormal situations in which large numbers of annunciators alarm is to "Turn off that _________ noise!" This often causes delays in responding to critical indicators.

15. During transient conditions, shifts to different power levels, shutdowns, startups, and other out-of-the ordinary conditions, the large number of annunciators that come on can easily mask safety-related annunciators because the operator is intent on coping with the unusual condition.

looks back. This error is made likely by the plethora of "normally on" lights, most of which do not require any operator action. He may then fail to take any action because the message he reads requires none, or he may remember that he has already taken care of the problem indicated by the incorrect legend light. The second reading error of consequence can occur because the legend light may be some distance away, and the operator may not walk over close enough to see the legend clearly. He sees the correct legend, but misreads it because of the distance.

Based on studies described in Chapter 19, we estimate that the probability of any reading error, including the two types above, is .001 (.0005 to .005) for an annunciated legend light. This estimated HEP is independent of the number of alarms and of operating conditions. Recovery factors will depend on the significance of the message the operator believes he has read, or on the occurrence of further signals.

In the case of automatic printout equipment, we estimate a zero reading error for the general sense of a message; e.g., the oil coolant temperature is high. If he reads the printout at all, he will read such messages correctly. For coded messages, or for series of numbers, the same reading errors described in Chapter 11 apply.

Scanning Errors for Unannunciated Conditions of Annunciated Displays

If an operator has failed to initiate action after canceling the auditory and blinking indications of a legend light, there is still some probability of his recovering from the error. Similarly, if the previous shift operator has turned off these indications without initiating action, there is some chance that the oncoming shift operator will see the signal and respond to it.

10-7 Scanning Errors for Unannunicated Conditions of Annunciated Displays

For any application, it has to be judged whether the typical control room operator scans the annunciator panels hourly or only when an alarm sounds or some other cue signals him. If there is no basis for such a judgment, as for example when the analysis is being done on plants in general, it is suggested that only one scan be assumed in the absence of alerting cues, and that that scan be assigned to the initial audit. Regardless of the frequency of scanning, the exponential decline in the probability of detection (Figure 9-1, p 9-7) is not applied to the detection of unannunciating legend lights on annunciator panels. Two opposing factors stabilize the probability of detection. First, all annunciated legend lights indicate abnormal plant conditions, so a certain minimal level of attention-getting is always associated with these indicators. Second, because of the large number of these legend lights that are usually on, the operator has to distinguish the new alarm from the background of old alarms. This involves an awareness of all the "acceptable" conditions reported by the old alarms, and is a very error-likely situation.

We estimate that the probability that an operator will detect a steady-on legend light (which requires action) on an annunciator panel is .05 per scan, except at the initial audit, for which we judge that the extra care doubles this probability to .1. Thus, the estimated probability of failure per hourly scan is 1 - .05 = .95 (.9 to .99) for all but the initial audit, when it is estimated to be .90 (.8 to .98). This error is judged to be independent of the number of steady-on indicators and of operating conditions.

Recovery factors for the failure to detect a steady-on legend light (which requires action) on the annunciator panel include the occurrence of functionally related signals and the onset of other annunciators. In the case of automatic printout equipment, we estimate a zero probability of recovery from later scanning in a shift. If the operator ignored the message when it alarmed, he would not be likely to check it later. Our rationale for this estimate is based on observation and on statements by operators that much of the information printed out is of little consequence and does not require immediate action.

Responses to Annunciating Indicators - Steady-State Operating Mode

It is assumed that responses to annunciating indicators under the steady-state operating mode will be made by the operator assigned to the control room. If two operators are assigned to the control room, human reliability may be increased. Human reliability will be decreased as additional annunciators compete for an operator's attention and as additional false alarms occur. At some point, of course, the number of annunciating indicators means that steady-state operating conditions no longer exist.

One Only Annunciating Indicator - One Operator

This section pertains to one of the only operator's responses to one of only one annunciating indicator or one of only one functional group of annunciating indicators when the plant is in the steady-state operating mode. The human performance model for one annunciator also applies to any group of annunciators that, by virtue of the common function they represent, will be regarded as a single unit by the operator; i.e., the annunciated indicators are completely dependent perceptually. A functional group might consist of two, three, four, or even five annunciators. No specific rules can be stated for defining such groups; the guiding principle is that the operator responds to them as if they were a unit. For reliability analysis, it is best to be conservative in defining such groupings.

Because of the compelling nature of the auditory alarm, failure to make a timely response to an annunciated signal is infrequent when there are no competing signals. However, in our interviews with operators and researchers associated with both U.S. and foreign plants, we established that occasionally an operator forgot an annunciated indication after the sound and blinking of the light were canceled. No data exist on which to base an estimate of the probability of this failure to respond in a timely manner. We assign a 10^{-4} (.00005 to .001) probability to this error as a best order-of-magnitude estimate. (For human responses, a 10^{-4} HEP is an extremely small number; it means that 9,999 out of 10,000 times the operator will initiate what he considers to be corrective action to a single annunciator within the allowed time.) A history of frequent false alarms could increase the 10^{-4} error estimate by one or more orders of magnitude. Although we have no data on the effects of false alarms from annunciator panels, data on the effect of false alarms in a different context (guard duty) indicate that they have a major influence.

The correct response is defined as perception of the alarm, acknowledgment of the alarm, decision as to what action is appropriate, and initiation of that action (which may include a decision to take no action). Note that the error term of 10^{-4} applies to the act of responding, not to the accuracy of the action taken, which could consist of operating switches in the control room, communicating to some other location for action to be taken at that site, and so on. The accuracy of the action taken (including the decision-making involved) must be evaluated separately.

Ordinarily, the operator acknowledges the alarm by turning off the audio signal almost immediately and then looking for the blinking light. When he finds which annunciator is blinking, he cancels the blinking function and reads the legend. The lamp will remain on until the problem has been corrected.

Depending on the nature of the alarm, the operator may be expected to initiate corrective action within some specific time. For safety-related alarms, this action would normally take place within a minute or so, although several minutes may be allowed in some cases. If corrective action is not initiated within, say, 1 minute after acknowledging the alarm, the probability that the action will be overlooked increases to 1.0 for some indeterminate period and then declines. The rationale is that ordinarily the operator will take action immediately after acknowledging an alarm. If he does not, it is because something more pressing requires his attention. While attending to the other event, he cannot attend to the initial event and so the HEP for the initial event becomes 1.0. Upon completing the required action on the second event, he is free to return to the initial event. However, he may have forgotten about the initial event while working on the second event, and the only indication remaining is the annunciator in the steady-on state, which is much less compelling than the blinking state.* Furthermore, this indication will normally be only one of several since there will be several steady-on indications (estimated at 20 or more) on the annunciator panels at any time. These constitute noise and must be filtered out by the operator.**

---

*This problem suggests an obvious human engineering design that we suggest be evaluated. This design would incorporate timing mechanisms which would cause steady-on annunciator indicators to resume blinking, perhaps with a unique auditory signal, if not corrected within some time period. Different annunciators might have different time periods for resumption of the blinking (and audio) signals. This concept is the basis of the snooze alarm found on some alarm clocks.

**This problem, too, is a candidate for an ergonomics solution, since several of these continuing indications may be related to repairs or other long-term conditions. One solution used in some plants is to paste repair stickers over such annunciator lights. However, a more effective solution would be to use translucent caps over these lights so that the illumination itself would be reduced. Reduction of the illumination would still enable an operator to see a blinking signal if the system were put back into service without removing the special cap.

A probability of 10^{-3} (.0001 to .01) is estimated for failure to respond to a steady-on annunciator within 1 minute after the interrupting task has been taken care of. The rationale is that the interruption increases the operator's error probability by a factor of 10 (from the basic 10^{-4} HEP), due to disruption of his short-term memory. However, if he does not initiate action within some very brief period (say 1 minute), we assume that he has forgotten the alarm, and the steady-on indication blends into the background of other steady-on indicators on the annunciator panels. The probability of his responding to it later will be much lower -- a .05 probability of detection per scan of the annunciator panel, assuming hourly or fewer scans per shift. The estimated Pr[F] is .95 (.9 to .99).

Figure 10-1 outlines the error probabilities for initiation of corrective action as a function of time after the initial annunciator comes on. We have arbitrarily assumed that, in the case of an interruption, the first scan takes place at the end of the minute following the hypothesized interruption.

The figure shows some of the considerations involved in accounting for time spent by operators as it relates to the probability of oversight in responding to an annunciator. The times involved vary. For example, some safety-related annunciators require a quicker response than do others. The 10^{-4} basic HEP is the estimate to use in answer to the question: What is the probability that an operator will fail to initiate corrective action (whether or not the action is the correct action) within the required time constraints, given a single alarmed annunciator (or one functional group of annunciators) and steady-state operating conditions without any interruptions)?

If there is some distraction or interruption before the operator can decide on a course of action, the initial HEP of 10^{-4} is increased by a factor of 10 to 10^{-3}. Figure 10-2 illustrates how to determine the estimated failure

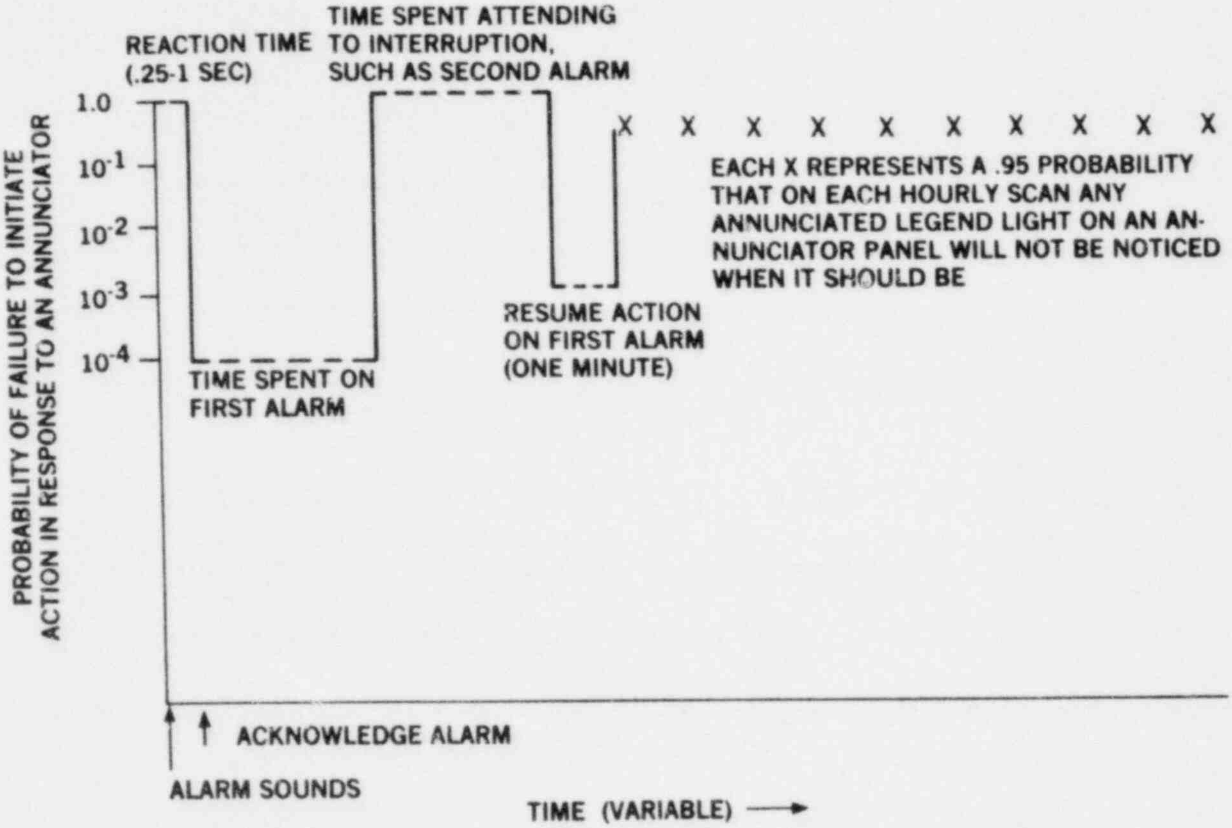
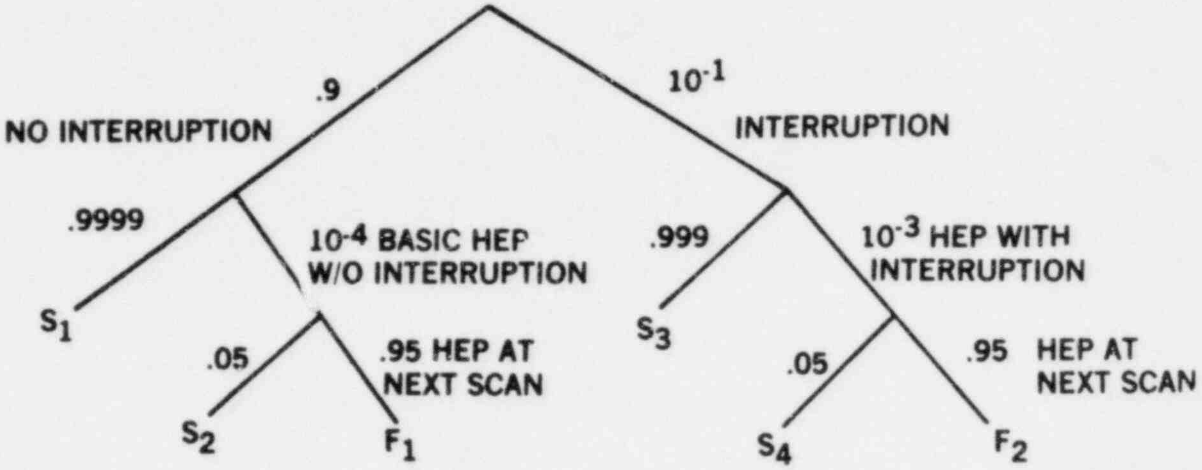


Figure 10-1. Initiation of Action in Response to Annunciators in Control Room, Given One Operator, Only One Annunciating Indicator, and Steady-State Operating Mode



$$\Pr[F] = F_1 + F_2$$

$$= (.9 \times 10^{-4} \times .95) + (10^{-1} \times 10^{-3} \times .95)$$

$$\approx 8.55 \times 10^{-5} + 9.5 \times 10^{-5}$$

$$\approx 18.05 \times 10^{-5}$$

$$\approx 2 \times 10^{-4}$$

Figure 10-2. Estimated Pr[F] to Initiate Corrective Action in Response to an Annunciator by the End of the Next Scan After Onset, Assuming Hourly Scans to take action to an annunciator by the end of the first scan following the onset of an annunciated indication, assuming there is some basis for estimating the percentage of times that an operator will be distracted or interrupted when responding to an annunciator. An HEP of .05 is estimated as the detection probability per scan for an annunciated legend light which had its blinking and auditory functions canceled and no further action taken in response to it. Assuming an interruption probability of 10%, the total failure probability is the sum of the failure paths F_1 and F_2, and is calculated as 2×10^{-4}. (The tree can be expanded to estimate the Pr[F] given additional scans.) The total probability of failure to initiate corrective action by the end of the first scan after the alarm is not greatly different from the basic 10^{-4} HEP. In a real-world situation, an interruption probability as high as 10% is unlikely under the steady-state operating mode. Therefore, in most cases, we can disregard the effects of interruptions and use the basic 10^{-4} probability for failure to initiate action in response to an annunciator, given there is only one annunciating indicator.

The overriding influence of the .9999 estimated probability of initiating action to an annunciator when it first comes on is illustrated by assuming that the annunciator illuminates after the fourth and before the fifth hourly scan and that the operator did not respond to it (except for the usual canceling of the auditory and blinking signals). He now has four hourly scans left in his shift, which means four chances to detect and respond to the legend light. Since the probability of his detecting the steady-on annunciator light is .05 per scan, his overall probability of detecting it by the end of the shift is:

$$\Pr[S] = 1 - .95^4 \approx .19$$

If hourly scans are not assumed, the next chance to see the steady-on legend light will occur during the initial audit of the next shift. Our model indicates that unless the operator responds appropriately (a probability of .9999) when the annunciator first comes on, he has a relatively low probability of recovering from this error on a timely basis.

One Annunciating Indicator - Two Operators

If two operators are assigned to the control room as active operators, we assume MD between them as per Equation 9-2, p 9-16. Thus, the basic error probability of 10^{-4} to acknowledge an annunciator and take corrective action would be modified per the equation for MD. If A is the HEP for person "A," and B|A is the CHEP for person "B" given that person "A" has failed, the total failure probability is:

$$\begin{aligned} \Pr[F|MD] &= A \times B|A \\ &= 10^{-4} \times .15 \approx 10^{-5} \end{aligned}$$

The above estimate is based on the assumptions that both operators are indeed present and that they have been instructed to monitor <u>all</u> the panels. If either assumption is not met, the above estimate will be too low. Seminara et al (1976) noted that when two operators were assigned to a reactor control room they both looked at all the panels. However, often only one operator is present in the control room. There are differences among plants and among shifts in a plant. If two operators are assigned, and if we assume that they will monitor the control boards simultaneously only half the time, the above equation would be modified as follows:

Pr[F] = [(failure probability for one person) x (percent of time first person only is available)] + [(failure probability for one person) x (percent of time second person is available) x (conditional probability of failure of second person assuming MD)] = $[10^{-4} \times .5] + [10^{-4} \times .5 \times .15] \approx 6 \times 10^{-5}$

Multiple Annunciating Indicators - One Operator

(In the following discussion, the term "one annunciator" also refers to a set of annunciators to which trained operators respond as a single unit.)

If an operator has to attend to two or more annunciating indicators, there is an increased load on him, and some decrease in his reliability is expected. As the number of annunciating indicators increases, the operator load will increase geometrically.

Operating personnel have two primary responsibilities: keeping the plant online and ensuring its safe operation. Because serious safety problems rarely occur at a plant, most of the operator's attention is directed to the instrumentation and controls related to the first responsibility. Interviews with operators indicate that they do not expect serious safety problems. Furthermore, they have confidence in the ability of their plant's safety systems to cope with possible safety problems automatically. These attitudes, coupled with the usual lack of practice in dealing with the unexpected, may result in reluctance on the part of an operator to take action that would interfere with keeping the plant online. To what extent the TMI accident and subsequent changes in operator training and onsite practice will affect these attitudes is not known.

It follows that if several annunciators sound simultaneously, an operator will probably give priority to those that are related to the maintenance of power output; i.e., the economic considerations. This is not meant to imply that the operator will purposely ignore safety-related annunciators in favor of annunciators related to keeping the plant online, but the tendency might be there. This tendency is strengthened as the number of false alarms in safety-related annunciators increases. In the EPRI Review some operators complained that false alarms were frequent. Estimates included "occasional," 15%, 30%, on up to 50% false alarms, 50 to 100 per shift, and even 100 per hour in one unusual case.

For many transients, the distinction between safety-related systems and economic-related systems is academic. Furthermore, in the case of loss of main feedwater, if backup auxiliary feedwater is not supplied on a timely basis, both safety and economic considerations are affected. Since loss of main feedwater will automatically trip the turbine and reactor, the above "conflict" between economic and safety considerations is not very important unless it affects subsequent operator actions.

There is still another consideration in attempting to predict the operator's response to some annunciator or group of annunciators. Often the operator will have standing orders in the event of a turbine/reactor trip. For example, at some plants where the transfer from main to auxiliary feedwater must be done manually, the operators are instructed to shift to auxiliary feedwater immediately <u>whenever</u> there is a turbine/reactor trip. The effect of such standing orders on operator response to related annunciators is difficult to estimate; a case study involving this situation is presented in Chapter 21.

Our model for multiple annunciating indicators does not address unique factors such as the above. It is based on the simplifying assumption that all annunciating indications are equal in importance and attention-getting. The user must modify the model for circumstances under which this assumption is invalid. Our auxiliary feedwater case study in Chapter 21 (p 21-14) shows how we handled this problem for one application.

The model is expressed in the following two equations and the resultant HEPs in Table 10-2:

$$\Pr[F_i] = \begin{Bmatrix} 10^{-4}, \ i = 1 \\ 2^{i-2} \times 10^{-3}, \ 1 < i \leq 10 \\ .25, \ i > 10 \end{Bmatrix} \qquad (10\text{-}1)$$

where $\Pr[F_i]$ is the failure to initiate action in response to the $\underline{i}^{\text{th}}$ annunciator (or completely dependent set of annunciators) in a group of $\underline{n}$ annunciating indicators.

$$\overline{\Pr[F_i]} = \sum_{1}^{\underline{n}} \frac{\Pr[F_i]}{\underline{n}} \qquad (10\text{-}2)$$

where $\overline{\Pr[F_i]}$ is the failure to initiate action in response to a randomly selected annunciator (or completely dependent set of annunciators) in a group of $\underline{n}$ annunciators. A lower uncertainty bound of HEP ÷ 10 and an upper uncertainty bound of HEP x 10 are assigned to each $\Pr[F_i]$ or $\overline{\Pr[F_i]}$, with an absolute lower bound of 5×10^{-5} and an absolute upper bound of .999.

The doubling of the estimated HEPs for annunciators after the second one (instead of a tenfold increase) is based on allowances for the arousal effect of the situation that causes a number of alarms to sound almost simultaneously. There is still a substantial increase in the probability that an operator will overlook some annunciators after he has canceled their sound and blinking because of the increase in task load as the number of competing

Table 10-2. Estimated Pr[F] for Multiple Annunciating Indicators

$Pr[F_i]$* for each Successive Annunciator (ANN) or Completely Dependent Set of ANNs

Number of ANNs	1	2	3	4	5	6	7	8	9	10	$\overline{Pr[F_i]}$**
1	10^{-4}										10^{-4}
2	10^{-4}	10^{-3}									6×10^{-4}
3	10^{-4}	10^{-3}	2×10^{-3}								10^{-3}
4	10^{-4}	10^{-3}	2×10^{-3}	4×10^{-3}							2×10^{-3}
5	10^{-4}	10^{-3}	2×10^{-3}	4×10^{-3}	8×10^{-3}						3×10^{-3}
6	10^{-4}	10^{-3}	2×10^{-3}	4×10^{-3}	8×10^{-3}	1.6×10^{-2}					5×10^{-3}
7	10^{-4}	10^{-3}	2×10^{-3}	4×10^{-3}	8×10^{-3}	1.6×10^{-2}	3.2×10^{-2}				9×10^{-3}
8	10^{-4}	10^{-3}	2×10^{-3}	4×10^{-3}	8×10^{-3}	1.6×10^{-2}	3.2×10^{-2}	6.4×10^{-2}			2×10^{-2}
9	10^{-4}	10^{-3}	2×10^{-3}	4×10^{-3}	8×10^{-3}	1.6×10^{-2}	3.2×10^{-2}	6.4×10^{-2}	1.3×10^{-1}		3×10^{-2}
10	10^{-4}	10^{-3}	2×10^{-3}	4×10^{-3}	8×10^{-3}	1.6×10^{-2}	3.2×10^{-2}	6.4×10^{-2}	1.3×10^{-1}	2.5×10^{-1}	5×10^{-2}
11-15	$Pr[F_i]$ for each additional ANN beyond 10 = .25										.10
16-20											.15
21-40											.20
>40											.25

---

[*] Estimated uncertainty bounds of HEP ÷ 10 and HEP x 10 are assigned to each $Pr[F_i]$ or $\overline{Pr[F_i]}$, with an absolute lower bound of 5×10^{-5} and an absolute upper bound of .999.

[**] $\overline{Pr[F_i]}$ is the expected Pr[F] to initiate action in response to a randomly selected ANN (or completely dependent set of ANNs) in a group of ANNs competing for the operator's attention. It is the arithmetic mean of the $Pr[F_i]$s in a row, with an upper limit of .25.

alarms increases. There may even be a point at which some operators may deliberately ignore the annunciators (except for canceling the disruptive sound), as evidenced by the TMI accident. The large upper bound is an attempt to include such incidents.

The cutoff HEP of .25 for the tenth annunciator (or 10 sets of completely dependent annunciators) was selected because more than 10 implies a stressful situation, such as a transient, and an error probability of .25 is assumed under high stress (p 17-17). Thus, for all annunciators beyond the tenth, an estimated HEP of .25 is assigned.

The $\overline{\Pr[F_i]}$ column in Table 10-2 is for use in reliability analyses in which the specific order of attending to annunciators cannot be predicted and the probability of failing to initiate action in response to a specific annunciator is of interest. In such a case, this probability is taken as the arithmetic mean of the $\Pr[F_i]$s of the sounding annunciators, because any annunciator has an equal chance of being first, second, or $\underline{n}^{th}$ to be selected. Since the values in Table 10-2 are speculative, the estimates for $\overline{\Pr[F_i]}$ for more than 10 annunciators are grouped as shown in the table; i.e., 11 to 15, 16 to 20, 21 to 40, and over 40 annunciators. Although $\overline{\Pr[F_i]}$ will never equal .25, this value is approached when $\underline{n}$ exceeds 40.

Multiple Annunciating Indicators - Two Operators

For the case of two operators, the error terms above should be modified as described in the section "One Annunciating Indicator - Two Operators," p 10-15.

One problem that has been mentioned in both the EPRI Review (Seminara et al, 1976) and the Kemeny (1979) report is that, as the number of annunciators sounding off increases, more and more personnel want "to get into the act," and a very confusing situation can arise, with errors in communication between people, and, on occasion, one operator taking actions that negate the actions of another operator. Of course, there are no means of predicting such effects - these are functions of plant discipline and administrative controls. However, for the steady-state operating situation the assumptions stated earlier regarding the interaction of two operators provide the most reasonable estimates of the advantages of two versus one operator.

For abnormal conditions in which more than two operators may be involved in the control room, see "Effects of Several Operators in an Abnormal Situation," p 17-22.

Responses to Annunciating Indicators - Maintenance or Calibration Operations

During certain calibration or maintenance procedures, the technicians may have to communicate with the control room operator by phone or intercom. For example, in calibrating setpoints that are alarmed in the control room, the technician will want to know whether the annunciator alarms at a certain level. In this type of activity, the control room operator is an active partner and is fully alert to annunciated indicators. For this reason, we judge that the probability of his responding to other annunciated indicators should not be degraded.

In other procedures, however, the control room operator may be passive in that alarms sound frequently and he merely notes that these alarms are part of the calibration procedure and turns off the audio and the blinking light indications. We have observed that some operators become annoyed by these constant interruptions and eventually may turn off the audio and blinking light without carefully checking to ensure that the annunciating indicator does indeed relate to the ongoing calibration or maintenance. Under these conditions, we assign an increase of an order of magnitude to the estimated HEP for steady-state operating conditions when a single annunciator comes on. Thus, for the one-operator situation, the estimated oversight probability would be .001 (.0001 to .01).

If two or more annunciators sound more or less simultaneously the operator should recognize an indication that is not associated with the calibration procedure. If the procedure involves just one annunciator, we judge that the onset of two annunciators will be perceived by the operator as "something different." This arousal effect should offset the interfering effects of the calibration procedure, and the model for normal operating conditions applies.

The above statements apply to annunciated legend lights. Ordinarily, annunciated printout equipment is not affected by calibration and maintenance as described above.

Responses to Annunciating Indicators - Anticipated Transient Events, Startup, and Shutdown

Anticipated transient events are pertubations in the steady-state operating condition of an NPP that may require rapid reactor shutdown.* Tables I 4-9 and I 4-12 in WASH-1400 list the PWR and BWR transients that were considered in the NRC reactor safety study. For many of these transients, the reactor will automatically trip, but for others (sometimes called Anticipated Transients Without Scram - ATWS) the operator will have to take action to control the situation. This may involve a rapid manual shutdown of the reactor. In any case, transient events are very demanding of an operator's attention.

---

*Per WASH-1400 (p I-58), in 150 reactor years no unanticipated transients have occurred, while there are about 10 anticipated transients per reactor year. For purposes of human reliability analysis, we regard unanticipated transients as psychologically similar to a small LOCA.

He will tend to concentrate all his attention on coping with the transient event and will be less likely to attend to annunciated signals not directly related to the transient.

Similarly, certain startup and shutdown procedures are also very demanding of an operator's attention, reducing the probability of his responding to a safety-related annunciator in time.

We estimate the same probabilities of failure to respond to annunciators related to transient events and to startup or shutdown procedures as were estimated for annunciators in general. However, when the plant is in a transient mode or in the shutdown or startup mode the estimated probability of the operator's failing to respond to an annunciated indicator not directly related to these conditions is increased by an order of magnitude. This increase applies also to the values in Table 10-2 for annunciated legend lights and to the probability of failure to respond to an annunciated printout.

Responses to Annunciating Indicators - Loss-of-Coolant Accidents

We are pessimistic about an operator's capabilities in coping with a serious emergency such as a large LOCA. When a large LOCA is simulated in a dynamic simulator, the noise and confusion are often overwhelming, even for experienced operators. Errors of oversight and commission are frequent. When a small LOCA is simulated, there is more time to take reasoned action, and errors are less frequent.

As yet we do not have a data bank of operator responses to LOCAs (simulated or actual), so estimates of human reliability in large or small LOCAs are highly conjectural. The HEPs listed in Table 10-2 allow for considerable degradation of operator performance under the stress of a large number of simultaneously sounding annunciators - we don't believe that these figures will degrade further under the stresses of a LOCA.

Given the high probabilities of effective automatic responses of ESFs to a small or large LOCA, the most important potential post-LOCA errors are those made by operators during manual switching from the injection mode to the recirculation mode or at later times during the recirculation mode. In Chapter 21 (p 21-1) we present the WASH-1400 analysis of changing from the injection to the recirculation mode and include changes to the original HEP estimates based on this handbook.

CHAPTER 11. UNANNUNCIATED DISPLAYS

As described in Chapter 10, annunciators alert the operator if some condition in the NPP is abnormal and requires immediate attention. However, most of the displays in an NPP are unannunciated. These are the displays that indicate the moment-by-moment status of selected plant parameters, and these are the ones the operator uses to run the plant. Whereas nearly all annunciated displays are legend-type displays that direct attention to some specific component or subsystem, the unannunciated displays are of a variety of types that present information with different degrees of precision. Unannunciated displays may be grouped in the following broad categories:

Meters

Digital displays

Chart recorders

Indicator lights (lamps or legend lights)

Graphs

This chapter describes some PSFs influencing the use of the above displays and presents HEP estimates related to the three major uses of displays: reading for quantitative information, check-reading for qualitative information, and periodic scanning of displays for abnormal indications or trends. Sources for derived HEPs in this chapter are described in Chapter 19. The scanning model presented in this chapter is based largely on our judgment, and can be modified when more data become available.

Major Performance Shaping Factors for Displays

In deriving estimates of BHEPs for unannunciated displays in the control room, the most relevant PSFs are the following:

(1) Stress level of the operator

(2) Rate at which the operator must process signals

(3) Frequency with which a particular display is scanned

(4) Relationship of the displays to annunciators or other attention-getting devices

(5) Extent to which the information needed for operator decisions and actions is displayed directly

(6) Human engineering related to the design and arrangement of the displays

In this chapter, the steady-state operating mode of the NPP is assumed for the first PSF, above, with the operators functioning at a low-to-optimal level of psychological and physiological stress (Figure 3-13, p 3-56). This range of stress is assumed for routine tasks of maintenance and calibration also. All HEPs in this chapter are premised on the optimal level of stress.

The second PSF relates to the first in that a requirement for rapid processing of information is often associated with a higher-than-usual level of stress. Best performance is usually obtained when the task is self-paced; i.e., the person proceeds at a comfortable pace. If the person has to function at a much higher rate, he is more error-prone. A requirement for processing signals at a high rate is a form of stress, time stress. Time stress does not affect performance in the same way as emotional stress; time stress usually leads to errors of omission, whereas emotional stress may result in much more severe incapacitation. A combination of time stress and emotional stress will usually result in the greatest level of incapacitation or the greatest number of inappropriate responses. In the steady-state operating condition of a plant, the stress level of the operator varies between low and optimal, and his required signal-processing rate is adequate since he has ample time to read and interpret the displays.

The third PSF, scanning frequency, is a function of the relative importance of the display as perceived by the operator, the context in which it is being read, and specified procedures.

The fourth PSF, relationship to attention-getting devices, refers to the phenomenon that, even though a particular deviant display is not annunciated, attention will probably be drawn to it if it is part of an annunciated subsystem.

The fifth PSF relates to the content of the information displayed. If the operator has no *direct* indication of certain functions, he will have to deduce the information from other displays; e.g., a problem in the TMI accident was the absence of a direct indication of emergency coolant flow from the *Auxiliary Feedwater System* (AFWS). Interpretation errors are more likely with this type of design, especially under stressful conditions.

The sixth PSF, human engineering related to the design and placement of displays, is highly variable in NPPs. Generally, there has been no systematic application of this technology to the designs of existing NPPs. We estimate that the HEPs listed in the handbook will be reduced by factors of 2 to 10 if the displays and controls are improved by the incorporation of standard human engineering concepts such as those described in Seminara, Eckert, et al, 1979. Table 11-1 lists some of the human engineering deficiencies in existing plants, based on the EPRI Review (Seminara et al, 1976) and our own observations.

For a complete reliability analysis, the influence of dependence must be considered, as described in Chapter 7, and errors of both omission and

Table 11-1. Human Engineering Deficiencies of Displays in Nuclear Power Plants

1. Displayed information that is primarily of actual status only; normal status is not indicated
2. Poorly designed scales, and scale numeral progressions that are difficult to interpret; e.g., one scale division equals 0.5 units
3. Parallax problems in relating pointers to scale markings and numerals on meters
4. Placement of meters and recorders above or below eye level, making the upper or lower segment of the scale difficult to read
5. Meters or recorders that can fail with the pointer reading in the normal operating band of the scale
6. Glare and reflections
7. Too many channels of information on chart recorders
8. Illegible pen tracings or symbols on chart recorders
9. No warning before a chart recorder pen runs out of ink
10. Use of chart recorders where meters or digital readouts would be more appropriate; e.g., where lags in data can result in wrong decisions
11. Functionally related displays that are widely separated physically
12. Inconsistent coding and labeling among displays
13. Mirror-imaging of control rooms
14. Lack of limit marks on meters used for check-reading
15. Meters not arranged with "normal" segments in the same relative positions (to facilitate check-reading)
16. Displays and arrangements do not present the operator with a mental image of what is going on

commission must be treated. This chapter deals only with errors of commission; most of the errors of omission described in Chapter 13, "Valving Operations," also apply to displays.

Reading and Recording of Quantitative Information from Displays

Under steady-state operating conditions, there won't be many deviant indications on displays other than those already tagged. Although we are primarily concerned with the probability of failure to detect a deviant indication on displays in the control room (a check-reading function), an operator sometimes has to read an exact value from a display. The following sections present HEPs and uncertainty bounds related to the reading and recording of quantitative information. The estimated Pr[F]s for these cases are collected in Table 11-2.

A given error of commission may or may not have any system-significance. For example, in reading a value from a digital readout, a calibration technician may erroneously read 1-2-3-5 instead of 1-2-3-4. In such a case, the error may not be important. On the other hand, if he erroneously reads 1-4-2-3, the error might have serious consequences. In performing a human reliability analysis, then, one must identify the important errors and estimate probabilities for them. As an approximation, we may assume that all possible reading errors are equally likely. Thus, in the above example, the technician has an equal likelihood of reading any given digit incorrectly. Therefore, if one is not interested in errors in the last digit of the four, the basic reading error for 4-digit digital readouts of .001, described below, can be multiplied by .75 to yield .00075 for the three digits of interest.

Table 11-2. Pr[F]s for Errors of Commission in Reading and Recording of Quantitative Information

Reading Task	HEP
Analog meter	.003 (.001 to .01)
Digital readout	.001 (.0005 to .005)
Chart recorder	.006 (.002 to .02)
Printing recorder with large number of parameters	.05 (.01 to .2)
Graphs	.01 (.005 to .05)
Values from indicator lamps that are used as quantitative displays	.001 (.0005 to .005)
Recognize that an instrument being read is jammed, if there are no indicators to alert the user	.1 (.02 to .2)

Number of Digits to be Recorded	HEP
≤ 3	Negligible
> 3	.001 (.0005 to .005)

(In practice, we would round this .00075 back to .001, showing that our data are too inexact to permit a distinction between HEPs in reading 3 or 4 digits.)

Meters and Digital Readouts

Based on data reported in Chapter 19, the estimated probability of a reading error is .003 (.001 to .01) for analog meters and .001 (.0005 to .005) for 4-digit digital readouts. If the readings are to be recorded manually, there is also an estimated error probability of .001 (.0005 to .005) in recording the reading if more than 3 digits are to be written down. For 3 digits or less, the recording error probability is judged to be negligible.

Chart Recorders

Most chart recorders are analog displays with the added feature of providing a record of the monitored parameters. Thus, they indicate the recent history of each parameter and enable the user to note any trends. The accuracy of chart recorders is somewhat less than that of well-designed panel meters of the same size (because of scale differences, pen lag, line width, etc), but this is of minor consequence. Ordinarily, we would assume an HEP for reading chart recorders that would be only slightly greater than the HEP for reading panel meters. However, in many NPPs, chart recorders are considerably more difficult to read than are comparable panel meters because of positioning, scaling, chart characteristics, pen characteristics, or multiplexing of data. Also, the pens of chart recorders are more apt to stick than the pointers of panel meters. The extent to which these disadvantages combine naturally varies; as a working estimate for reading chart recorders we suggest doubling the estimated HEP of .003 for reading meters to .006 (.002 to .02).

In addition to pen-writing chart recorders, there are printing recorders that periodically stamp a number on the chart. This number corresponds to the monitored parameter, and the position where it is printed corresponds to the value of the parameter. A large number of parameters can be recorded on a single chart. Recorders of this type are particularly susceptible to reading errors because of faulty registration, faulty printing, double printing, and differences in scale associated with the different parameters. Based on these difficulties, the HEP for reading recorders of this type is estimated to be .05 (.01 to .2). Most of these errors will be due to confusing the identifying numbers (or letters) of the parameters.

Graphs

Graphs are not used very much in NPPs. When they are, the operators freely use any aids they desire, such as rulers, pencils, etc. It is difficult to read a graph with precision, but it is unusual for graphs to be used where precise interpolation is required. The estimated HEP for reading graphs is .01 (.005 to .05).

Indicator Lamps

Occasionally, a set of status lamps is used as a quantitative display. For example, in one plant the containment sump level is indicated by five lamps labeled 4, 18, 48, 64, and 205 inches (see Figure 3-10, p 3-39). As water in the sump rises, the lamps light in sequence to indicate the water level. Under stress-free operating conditions, this type of indication should be relatively free of human error even though the design violates a populational stereotype (p 3-38). We assess the reading error to be the same as that for reading a digital readout, .001 (.0005 to .005). (Because it does violate a populational stereotype, for the reading HEP under stressful conditions we multiply the BHEP of .001 by 10.)

Some Recovery Factors for Errors in Quantitative Readings

Errors in reading displays can be recovered in several ways. Two of the most common are discussed in the following two sections. The first relates to the operator's expectancy of what the reading should be. The second is the recovery obtained when two people perform a reading task; i.e., the use of human redundancy.

Recovery Factors Based on Operator Expectancy

Note that the HEPs presented above apply to the probability that the error will occur at all -- they do not consider the operation of recovery factors that can alert the user to his error. For example, the HEP for recording 4 digits is .001; this includes the most common error, that of transposition -- a person reads 3821 and writes 8321. If the entry is to be used immediately, as when calibrating an item, the resulting figure is so obviously deviant than usually it would be questioned and corrected. Similarly, transposing the two middle digits results in a grossly deviant figure. When we arrive at the last two digits, the error may or may not be obvious, depending upon the significance attached to them. If they are as significant as the others, again the error usually will be noticed. If not, the error itself may not matter, as when a data form requires more accuracy than the situation warrants. The important concept about errors of the type discussed in this section is that gross errors in taking data from displays usually will be noticed. However, in preliminary reliability analyses, we often take the conservative position that recovery factors will not operate when needed. To avoid undue pessimism, recovery factors must be included in a final analysis.

The Use of Human Redundancy in Quantitative Readings

If two people act as a team in reading and recording display indications, the error probability will be a function of the procedures followed by the team. For example, if one operator reads the indication aloud, the other operator records it, and then they both check each other, the benefits of human redundancy are maximal.

Consider the opposite case, when the team members do not check each other -- for example, if one operator reads the indication aloud and the other records it without any checking. In this situation we would not allow for any benefits of human redundancy. It is possible, of course, that if the reader makes a gross error, the person doing the recording might notice it. However, this type of team interaction is so passive that we assess the probability of a wrong entry as the sum of the reading and recording errors. (There also is a possibility of an error of communication between the team members, but this will be disregarded here).

The highest reliability would be attained if the two people read and recorded individually on separate sheets. In such a case, complete independence of the two team members might be assumed, and the error probability for an individual would be squared if there were some error-free comparison of the two records. Team members tend to have an alerting effect on each other in actual practice, and usually there is some informal checking on each other even if not required by written procedures, so that the performance of teams will usually be more reliable than that of individuals. Often, the reliability increases because a team is more likely to follow plant policies than is an individual working alone. We are unable to estimate the quantitative value of this informal recovery factor because it is highly variable. To be conservative, we assign no recovery factor for this effect. Further discussion of human redundancy appears in Chapter 15.

"Funneling" of Attention in Use of Displays

In responding to a display, an operator may focus his attention on a particular display to the exclusion of all others. Often, the operator will initiate the indicated action and concentrate his attention on that display, waiting for a change in the readout. This "funneling" of attention is more likely to occur when the operator is under stress.

An occasional display malfunction is "sticking"; i.e., a pointer on a panel meter or a pen on a chart recorder jams for some reason and no longer yields useful information. Usually there are several redundant displays for any significant parameter, and the operator can refer to one of them for the required reading until the primary display is repaired. However, there is a strong tendency to focus on just one display without cross-checking. Because of the operator's involvement in the corrective action to be taken, this is most likely to occur when the display sticks in a position indicating the need for immediate corrective action. It is less likely to occur when the sticking display does not indicate a need for immediate action because the operator will be scanning the associated displays as well. When an operator uses an instrument that has jammed without any indication to that effect, we estimate a probability of .1 (.02 to .2) that the operator will fail to cross-check until some other indication, such as an alarm, alerts him that something is amiss.

Check-Reading of Displays

In many cases, displays are merely "checked" rather than read quantitatively. That is, the operator refers to the display to determine that the monitored parameter is within certain limits, rather than to determine the exact value of the reading. The check-reading may be merely to note go/no-go indications such as which indicator light is on (e.g., is it a red light, or is it a green light) or whether the pointer is still within the acceptable range on a meter. At other times, the check-reading may require more detailed qualitative discrimination; e.g., has the pointer moved upwards since the last time it was checked? The following sections describe check-reading of meters, digital readouts, chart recorders, and indicator lights. The estimated HEPs associated with these tasks are in Table 11-3. The HEP estimates for check-reading apply to displays that are checked individually for some specific purpose, such as a scheduled requirement, or in response to some developing situation involving that display. For this reason, these HEPs are much smaller than the ones related to the more passive periodic scanning of the control boards, discussed later.

Check-Reading of Meters and Chart Recorders

Check-reading of meters and chart recorders is facilitated by the use of limit marks to indicate the limits of acceptable readings. The "red lines" on tachometers are a familiar example.

The estimated BHEP of commission for check-reading meters is .003 (.001 to .01). This applies to meters without limit marks. If there are easily visible limit marks, we estimate that the error probability is reduced by a factor of 3; i.e., to .001 (.0005 to .005). For analog-type chart recorders, the above HEPS are doubled to .006 (.002 to .02) for charts without limit marks and to .002 (.001 to .01) for charts with limit marks.

In most NPPs, the meters used for check-reading are purchased and installed without limit marks. However, the operators usually add informal limit marks in the form of tape, grease-pencil lines, etc., that are almost as

Table 11-3. HEPs for Check-Reading of Displays

Check-Reading Task	HEP
Digital indicators (these must be read - there is no true check-reading function for digital displays)	.001 (.0005 to .005)
Analog meters with easily seen limit marks	.001 (.0005 to .005)
Analog meters with difficult-to-see limit marks, such as scribe lines	.002 (.001 to .01)
Analog meters without limit marks	.003 (.001 to .01)
Analog-type chart recorders with limits	.002 (.001 to .01)
Analog-type chart recorders without limit marks	.006 (.002 to .02)
Confirming a status change on a status lamp	Negligible
Checking the wrong indicator lamp (in an array of lamps)	.003 (.001 to .01)
Misinterpreting the indication on the indicator lamps	.001 (.0005 to .005)

effective as factory-printed limit marks. At some plants, management policies do not allow the use of limit marks on meters, and the operators resort to the use of "invisible" limit marks, such as fine scribe lines. Since these are not easily seen, we assign an HEP between the two values above, .002 (.001 to .01).

Digital Readouts

With analog displays, check-reading does not involve a quantitative reading -- the indication is either "in" or "out." The relative position of the pointer provides all the required information. With digital displays, there are no such positional cues -- the display must be read, so we judge that the check-reading HEP of commission for digital displays is the same as for their quantitative readings -- .001 (.0005 to .005).

Check-Reading of Indicator Lights

Indicator lights are used to indicate the state of some component or subsystem. The discussion that follows assumes either transilluminated legend lights or a plain lamp with a colored cover and a label above or below the lamp.

In certain applications in NPPs, color conventions are observed. For example, valve states are indicated by red for open and green for closed. In a few cases entire subsystems are arranged to comply with a "green-board" philosophy; that is, if all components in a subsystem are in the normal operating mode, all the indicator lights will be green.

Aside from those described above, few conventions are followed in the color-coding of indicator lights in NPPs, and the lamps do not indicate the normal operating states of the items they are monitoring. (In this context "normal" refers to when the plant is in the steady-state operating mode.)

The three usual cases in which an operator will observe a status lamp are when

1. Confirming a status change after a manual operation such as changing a valve state
2. Determining the status of a specific item (or group of items) for some immediate purpose
3. Conducting a survey, as in the course of an initial audit or at prescribed periods thereafter

The HEPs associated with the first two cases are as follows:

1. Confirming status change after an operation, such as changing the status of a MOV, is an active task. The operator has initiated the change and ordinarily will watch the indicator lights for confirmation of the response. Under normal operating conditions, the probability of his failing to note the status change is negligibly small and will be disregarded.
2. Checking the status of a specific indicator light (or group of lights) for some specific purpose is also an active task. Given the usual large number of similar indicator lights, we estimate a .003 (.001 to .01) probability of checking the wrong indicator light. Assuming that the correct indicator is addressed, there is some small probability (estimated as .001 (.0005 to .005)), that the indication will be misinterpreted (e.g., a reversal error will be made when looking at a pair of lamps consisting of a red and a green lamp).

The third case is discussed in the next section.

Detection of Deviant Unannunciated Displays During Periodic Scanning

The above sections describe the cases in which some specific cue directs an operator to look at a particular unannunicated display. For example, an annunciator alarms, and the operator checks meters functionally related to that annunciator. Or the operator follows some schedule that requires him to note some quantitative reading or to check-read some display. In such cases, there is a high probability of his detecting a deviant indication on the display he observes. The remaining sections in this chapter deal with the probability that an operator will detect some deviant indication on a display when no such cues are present. In such cases, detection will depend largely on his scanning pattern and frequency.

Hourly scans are assumed as the average practice of control room operators for those displays on which important information may frequently change. For other types of displays, unless one has information to the contrary, only one scan per shift is assumed, at the initial audit. For example, essentially static displays such as a meter showing the level of refueling water in a tank are not likely to be scanned more than once per shift. Some operators state that there are many displays they never check unless alerted by some signal or an alarm; e.g., status lamps for blocking valves.

The estimates below are based on the observation that, although he may "look at" a given display, there is a high probability that a deviant indication will not register on the consciousness of the operator unless some alerting cue is present. This error is classified as an error of commission since the deviation is "seen" but not perceived.

Deviant indications may appear on any of the dynamic displays described earlier in this chapter. The modeling differentiates between go/no-go displays, such as status lamps and legend lights, and displays that present additional information, such as meters, chart recorders, and digital readouts.

Status Lamps and Legend Lights

Without any alerting cues, the probability of detecting a deviant status lamp or legend light during a scan, when there is only one such deviant indicator, is very low. For scanning of a status lamp (e.g., noting that the green lamp is on when the red lamp should be on), our best order-of-magnitude estimate of this detection probability is .01 per scan; i.e., an error estimate of .99. For a legend lamp, the estimate of detection probability is doubled since the latter has more information to aid the operator. (The above and following HEPs and discussion also apply to pairs of lamps or legend lights which monitor the same function; e.g., the two status lamps at TMI which monitor the two blocking valves for the AFWS.)

In most applications, either one scan per shift or no scan is assumed for status lamps and legend lights. If more than one scan is made per shift, a constant detection rate for each scan is assumed throughout the shift, beginning with the initial audit. Thus, the estimated failure probabilities per scan to detect a given deviant status lamp or legend light, when there is only one deviant indicator, are, respectively, .99 (.98 to .998) and .98 (.96 to .996). (The exponential curve described in Chapter 9 is not applied since the estimated probabilities of detection are so low relative to meters and other dynamic displays for which the curve is assumed.)

Since the scanning of these types of displays is passive, and since it is possible that some displays are not checked at all in the absence of some alerting cue, we assume that if a deviant status lamp or legend light has not been detected in 30 days, it will not be detected in any further period. (This same 30-day cutoff is assumed for all periodic checks and scans.)

Analog Displays

For analog displays, we assume the exponential decrease in detection efficiency described in Chapter 9. For multiple deviant analog displays, and for more than one operator, special performance models are used. The next major section presents some numerical values to illustrate the special application of the models to meters and chart recorders and presents a rationale for the probability estimates given.

Digital Readouts

The scanning model for analog displays does not apply to digital readouts because the latter must be read rather than scanned. If a particular digital readout is read periodically, we assume that its criterion value (or range) is commonly known (e.g., reactor temperature) or that it is posted next to the display. In such a case, the only significant error is the error of reading, as discussed earlier.

Multiple Deviant Displays Present During Scanning

If more than one display become deviant during a shift and none are unannunciated, the estimated probability of detecting at least one of them is calculated using Equation 9-1 (p 9-13) or Table 9-1 (p 9-14). For two operators, use Equation 9-2, p 9-16. If the operator has detected one deviant display during scanning, the estimated probability of detecting functionally related deviant indicators follows the model for annunciators. That is, under normal plant conditions, detection of any deviant display should cause arousal such that related deviant indications take on the same attention-getting values as annunciated displays (see Table 10-2, p 10-19).

If there are any deviant displays not functionally related to the detected deviant display, it is assumed that the arousal effect will raise detection effectiveness to the initial audit level of effectiveness.

Detection of Deviant Analog Displays During Periodic Scanning

The scanning model in this section is based on hourly scanning of meters with limit marks, with no cue or special instruction that would draw attention to a particular meter. If there is some cue or special instruction, the HEPs for reading or check-reading should be used instead of the scanning model. The model is based on the assumption of only one deviant display and one operator. For more than one deviant display, and for more than one operator, see Equations 9-1 and 9-2 (pp 9-15 and 9-16).

The model is based on meters with limit marks, and adjustment is made for application to meters without limit marks or to analog-type chart recorders with and without limit marks. Scanning is a special kind of check-reading, and the modifying factors for check-reading errors may be applied to scanning errors when comparing different analog displays. The following sections present the scanning models and estimated HEPs for meters and chart recorders with and without limit marks.

Scanning of Meters with Limit Marks

The probability of an operator's detecting one of only one deviant display is assumed to follow an exponential decrease in detection efficiency over the shift if there are no alerting cues. The end spurt in detection ability is assigned to the initial audit in the following shift. We assume that the operator scans displays at hourly intervals, beginning with the initial audit and ending with the eighth scan an hour before the shift ends. Zero duration time for each scan is assumed for mathematical convenience (in typical reliability analyses the system failure estimates are not materially changed by this assumption). Thus, the inital audit, T_1, takes place at time zero, the second scan, T_2, at the start of the start second hour on the shift, and so on to the last scan T_8, which takes place at the start of the eighth hour on the shift. As noted on p 9-6, the initial audit in the subsequent shift is assumed to include any last-minute scan on the previous shift.

While the exponential shape of the curve is in general agreement with what is known about detection efficiency over time, a search of the literature revealed no data with which to assign probability estimates to the eight points on the curve. For example, Murrell (1969, pp 62-63) notes the lack of data on control room errors in large process plants. On the basis of our experience, we hypothesize a .95 probability of successful detection of one of only one deviant display at the initial audit, T_1, and a .1 probability of successful detection for the last hourly scan, T_8, given nondetection on the previous trials. With values assigned to these two points, the other points were determined mathematically (based on the exponential function) and are listed in the first column of Table 11-4. These figures represent the hypothesized hourly decline in detection probability throughout the 8-hour shift.

Equipment malfunctions are assumed to occur midway into the shift (Chapter 6). This assumption means that a meter indication related to some deviant condition in the plant will occur just before T_5. As explained on p 9-8, the simplifying assumption is made that the tabled detection probabilities hold regardless of when the deviation occurred. The estimates

Table 11-4. Estimated Per Scan and Cumulative Probabilities of Detection for Each Hourly Scan for One (or One Completely Dependent Set) of One Deviant Unannunciated Meter, With Limit Marks (p 1 of 2)

Trial (Scan) Number	$\Pr[S_i]$	$\Pr[F_i]$	$\Pr[S_{(i)}]$	$\Pr[S_{\leq i}]$
T_1	.95	.05	.95	.95
T_2	.69	.31	.03	.98
T_3	.50	.50	.008	.988
T_4	.36	.64	.003	.991
T_5	.26	.74	.001	.992
T_6	.19	.81	.0007	.9927
T_7	.14	.86	.0004	.9931
T_8	.10	.90	.0003	.9934

NOTES:

1. These estimates are rounded values. Four significant figures are used merely for completeness; in practice these should be rounded further.
2. $\Pr[S_i]$ is the probability of detection on the i^{th} trial, given nondetection on all previous trials either because of oversight or because there was no deviant condition.
3. $\Pr[F_i] = 1 - \Pr[S_i]$
4. $\Pr[S_{(i)}]$ is the probability of first detection on the i^{th} trial only, given that the deviant condition occurred before T_1 of the present shift but after T_8 of the preceding shift. For example, $\Pr[S_{(4)}] = .05 \times .31 \times .50$ (all from the $\Pr[F_{(i)}]$ column) $\times .36$ (from the $\Pr[S_i]$ column) $= .00279 \approx .003$.
5. $\Pr[S_{\leq i}]$ is the probability that detection occurs on or before the i^{th} trial, given that the deviant condition occurred before T_1 of the present shift but after T_8 of the preceding shift.

6. To calculate $Pr[S_{(i)}]$ and $Pr[S_{\leq i}]$ for deviant indications that occurred just before hourly scans other that T_1, start the calculations from the T_i of interest. (See the appendix to this chapter for some calculations.)
7. For HEPs < .5, the lower uncertainty bound is calculated as HEP x 0.2 and the upper bound is HEP x 2. For HEPs $\geq$ .5, the lower uncertainty bound is 1 - 2(1 - HEP) and the upper bound is 1 - 0.2(1 - HEP).
8. If fewer than eight scans per shift are assumed for a specific application, use the T_i values from the table according to the times for which the scans are assumed. Thus, if scans at the beginning and midway through the shift are the two scans assumed, use the $Pr[S_i]$ and $Pr[F_i]$ values for T_1 and T_5.
9. It is assumed that if the deviant meter has not been detected within 30 days, it will not be detected unless some other stimulus calls its deviation to the operator's attention.

in Table 11-4 refer to either a single display or a set of displays for which the probabilities of detection are completely dependent. The tabled values apply to the steady-state operating mode only and should not be applied to conditions in which disabling levels of stress can occur.

The estimates in Table 11-4 are predicated on the passive nature of scanning in comparison to check-reading or quantitative reading activities; the operator is looking around the control room to see "if everything is OK." Under steady-state operating conditions, he expects to find everything within proper limits, and usually they will be. He is not actively "probing" each indicator. He is inclined to accept readings as being within proper limits unless his attention is caught by a grossly deviant indication.

In unavailability calculations the mean number of trials (i.e., scans) to detection, $\bar{t}$, or the median number, M, are often of interest. Table 11-5 lists the mean and median numbers of trials to detection for various starting points. For example, the mean value of 3.33 and the median value of 4 represent the number of trials to detection for some unannunciated display that becomes deviant after T_4 and before T_5; i.e., midway into the shift.

To calculate $\bar{t}$ we base the calculations on a total of eight trials, since additional trials make increasingly small contributions to $\bar{t}$. To calculate $\bar{t}$, designate the first trial after the deviation as T_i, look up the values for that T_i in the first two columns in Table 11-4, and designate these as the probabilities of success and failure for the first trial of the calculation series. The calculations to be followed thereafter are illustrated in the appendix to this chapter.

To calculate M from any specified trial, start with the appropriate T_i in Table 11-4 and calculate the successive $\Pr[S_{\leq i}]$ values until reaching the trial in which the cumulative probability first equals or exceeds .50. A sample calculation is shown in the appendix to this chapter.

Table 11-5. Mean and Median Numbers of Trials to Detection for Deviations of Unannunciated Meters with Limit Marks that Occur Prior to Any Given Scan T_i from Table 11-2.*

Scan T_i	Mean**	Median
T_1	1.06	1
T_2	1.78	1
T_3	2.55	1
T_4	3.13	2
T_5	3.33	4
T_6	3.16	4
T_7	2.68	3
T_8	1.96	2

---

*The derivations of the mean and median values for T_5 are shown in the appendix to this chapter.

**The mean values are based on a total of 8 trials from any T_i of interest because, with the values given in Table 11-2, use of more than 8 trials does not change the mean materially.

Scanning of Other Types of Analog Displays

Table 11-6 presents the estimated probabilities of success and failure for each hourly scan for meters without limit marks and for chart recorders with and without limit marks. The $\Pr[F_i]$ values for T_1 were determined by multiplying the equivalent $\Pr[F_1]$ of .05 for meters with limit marks (Table 11-4) by a factor of 3 for meters without limit marks, a factor of 2 for chart recorders with limit marks, and a factor of 6 for chart recorders without limit marks. The $\Pr[F_i]$ values for T_8 were also based on the equivalent $\Pr[F_1]$ of .9 for meters with limits marks, as follows:

$$1 - \frac{1 - T_8 \text{ value of } \Pr[F_i] \text{ for meters with limit marks}}{\text{the appropriate factor of 2, 3, or 6}}$$

The values for T_2 through T_7 are based on the exponential relationship used in Table 11-4.

The appendix to this chapter shows how mean and median numbers of trials to detection of these types of unannunciated displays could be calculated for any given scan T_i from Table 11-6.

Table 11-6. Estimated Per Scan Probabilities of Detection for Each Hourly Scan for One (or One Completely Dependent Set) of One Deviant Unannunciated Display

	Type of Display					
	Meters		Chart Recorders			
	Without		With		Without	
Trial (Scan) Number	Limit Marks		Limit Marks		Limit Marks	
	$Pr[S_i]$	$Pr[F_i]$	$Pr[S_i]$	$Pr[F_i]$	$Pr[S_i]$	$Pr[F_i]$
T_1	.85	.15	.90	.10	.70	.30
T_2	.53	.47	.60	.40	.42	.58
T_3	.33	.67	.39	.61	.25	.75
T_4	.20	.80	.26	.74	.15	.85
T_5	.13	.87	.17	.83	.09	.91
T_6	.08	.92	.11	.89	.06	.94
T_7	.05	.95	.08	.92	.03	.97
T_8	.03	.97	.05	.95	.02	.98

NOTES:

1. $Pr[S_i]$ is the probability of detection on the i^{th} trial, given nondetection on previous trials either because of oversight or because there was no deviant condition.
2. $Pr[F_i] = 1 - Pr[S_i]$.
3. For HEPs < .5, the lower uncertainty bound is calculated as HEP x 0.2 and the upper bound is HEP x 2. For HEPs $\geq$ .5, the lower uncertainty bound is 1 - 2(1 - HEP) and the upper bound is 1 - 0.2(1 - HEP).
4. If fewer than eight scans per shift are assumed for a specific application, use the T_i values from the table according to the times for which the scans are assumed. Thus, if scans at the beginning and midway through the shift are the two scans assumed, use the $Pr[S_i]$ and $Pr[F_i]$ values for T_1 and T_5.
5. A 30-day cutoff is assumed as per note 9 in Table 11-4, p 11-22.

APPENDIX TO CHAPTER 11. CALCULATIONS OF MEAN AND MEDIAN NUMBERS OF TRIALS TO DETECTION GIVEN IN TABLE 11-5

Table 11-5 (p 11-24) lists the mean and median numbers of trials to detection of a deviant unannunciated meter with limit marks that becomes deviant prior to any given scan, T_i. The HEPs used in the derivation of these numbers are from Table 11-4 (p 11-21). For the unavailability calculations that are likely to be made, the values in Table 11-5 should be sufficient. However, since one purpose of the handbook is to introduce those not familiar with unavailability calculations to the methods used, these calculations are presented here in detail. The same types of calculations can also be used to derive mean and median numbers of trials to detection for displays other than meters with limit marks. That is, these calculations can be applied to the values in Table 11-6 (p 11-26).

The calculations below illustrate how the mean number of trials to detection, 3.33, was calculated for the case in which a deviant display occurs midway into the shift; i.e., just prior to T_5 in Table 11-5. As noted, we designate this T_i, the first inspection following the occurrence of the deviation, as the starting point for the iterative procedure used to calculate $\bar{t}$, mean trials to detection, and use the appropriate values from the first two columns in Table 11-4 in the following equation:

$$\begin{aligned}\bar{t} = {} & 1\Pr[S_i] + 2\Pr[F_i]\Pr[S_{i+1}] + 3\Pr[F_i]\Pr[F_{i+1}]\Pr[S_{i+2}] \\ & + \ldots + (8 - i + 1)\Pr[F_i]\Pr[F_{i+1}] \times \ldots \times \Pr[F_7]\Pr[S_8] \\ & + (8 - i + 2)\Pr[F_i]\Pr[F_{i+1}] \times \ldots \times \Pr[F_8]\Pr[S_1] \\ & + (8 - i + 3)\Pr[F_i]\Pr[F_{i+1}] \times \ldots \times \Pr[F_8]\Pr[F_1]]\Pr[S_2] \\ & + \ldots + 8\Pr[F_i]\Pr[F_{i+1}] \times \ldots \times \Pr[F_8]\Pr[F_1] \\ & \qquad \times \ldots \times \Pr[F_{i-2}]\Pr[S_{i-1}]\end{aligned}$$

where $\bar{t}$ is the mean number of trials to detection, and, after failure to detect on the last trial in the shift (T_8 in Table 11-4), the new shift takes over (T_1 in Table 11-4). The mean is based on 8 trials only beause the use of more than 8 trials does not change the mean materially.

To illustrate the application of this equation, $\bar{t}$ is calculated below:

$$
\begin{aligned}
\bar{t} = {} & 1\Pr[S_5] + 2\Pr[F_5]\Pr[S_6] + 3\Pr[F_5]\Pr[F_6]\Pr[S_7] \\
& + 4\Pr[F_5]\Pr[F_6]\Pr[F_7]\Pr[S_8] + 5\Pr[F_5]\Pr[F_6]\Pr[F_7]\Pr[F_8]\Pr[S_1] \\
& + 6\Pr[F_5]\Pr[F_6]\Pr[F_7]\Pr[F_8]\Pr[F_1]\Pr[S_2] \\
& + 7\Pr[F_5]\Pr[F_6]\Pr[F_7]\Pr[F_8]\Pr[F_1]\Pr[F_2]\Pr[S_3] \\
& + 8\Pr[F_5]\Pr[F_6]\Pr[F_7]\Pr[F_8]\Pr[F_1]\Pr[F_2]\Pr[F_3]\Pr[S_4] \\
= {} & 1(.26) + 2(.74)(.19) + 3(.74)(.81)(.14) \\
& + 4(.74)(.81)(.86)(.10) + 5(.74)(.81)(.86)(.90)(.95) \\
& + 6(.74)(.81)(.86)(.90)(.05)(.69) \\
& + 7(.74)(.81)(.86)(.90)(.05)(.31)(.50) \\
& + 8(.74)(.81)(.86)(.90)(.05)(.31)(.50)(.36)
\end{aligned}
$$

= 3.33 mean trials to detection given that the deviant display occurred midway into the shift.

To calculate the median, start with the $\Pr[S_i]$ for the T_i immediately following the time in which the deviant display is postulated to have occurred, and calculate the successive $\Pr[S_{\leq i}]$ values until reaching the trial in which the cumulative probability first equals or exceeds .50. Thus,

$$M = r$$

Where M = median number of trials to detection

and r = the smallest number of trials that satisfies

$$\Pr[S_{\leq i}] \geq .50$$

and $\Pr[S_{\leq i}]$ = the probability that detection occurs on or before the i^{th} trial, given that the deviant condition occurred before the T_i of interest but after T_{i-1}.

If, for example, we assume that the deviation occurred after T_8 on a shift but before T_1 of the next shift, the appropriate T_i is T_1. From Table 11-4 we see that the median number of trials to detection is 1 because the $\Pr[S_{\leq i}]$ value for T_1 already exceeds .50. If the deviant display occurs in the middle of the shift; i.e., just before T_5, the median number of trials is 4 because at the fourth term in the following equation the cumulative probability first exceeds or equals .50. That is, r = 4 because

$$\begin{aligned}
&\Pr[S_5] + \Pr[F_5]\Pr[S_6] + \Pr[F_5]\Pr[F_6]\Pr[S_7] \\
&+ \Pr[F_5]\Pr[F_6]\Pr[F_7]\Pr[S_8] \\
= \; &.26 + (.74 \times .19) + (.74 \times .81 \times .14) \\
&+ (.74 \times .81 \times .86 \times .10) \\
\approx \; &.54 \geq .50
\end{aligned}$$

Thus, if a deviant indication occurs just before T_5, 50% of the time it will be detected on or before T_8 (the end of the shift). As a matter of interest, the cumulative probability of detection $\Pr[S_{\leq i}]$ jumps to about .98 that a deviant display that was not detected on T_8 will be detected on T_1 of the next shift. Each further trial adds a small increment to the cumulative detection probability, until the next T_1, when its $\Pr[S_{\leq i}]$ increases from .9988 to .99997. Obviously, the cumulative probability never reaches 1.0, but approaches it as a near asymptote.

CHAPTER 12. MANUAL CONTROLS

Operation of an NPP involves hundreds of controls (e.g., switches) of many different types. Many of the switches in a plant control components or functions that are also automatically controlled in that the components or functions respond to signals from sensors or to computer commands. Manual controls are those handled by the operator; they are the means by which the human enters his inputs to the system. This chapter primarily addresses manual controls in the reactor control room. The controls outside the control room may be regarded as special cases requiring individual evaluation.

Types of Manual Controls

Almost all the manual controls in an NPP control room are electrical. These controls may be either continuous or discrete. Continuous manual controls do not have detents; they may be adjusted to any point within their range (e.g., a potentiometer). Discrete (or discontinuous) manual controls, which have detents, are used to select one of a limited number of states (e.g., a switch). Most of the controls in an NPP are discrete. The most common controls in NPPs are the following:

Multiposition selector switches

Transilluminated switches

Toggle switches

Valve-control switches

Handwheels

Pushbuttons

Rotary knobs

Levers

Cranks

Thumbwheels

Connectors (cables, jumpers, and interlocks)

Tools are also controls, but will be ignored here.

Errors in Manual Control Activation

When an operator reaches for a control, the decision to manipulate that control has already been made. This section does not address possible errors in arriving at that decision. Assuming that the decision was correct, we find the possibility for three errors in the manipulation of the control: (1) selection of the wrong control, (2) incorrect operation of the control, and (3) inadvertent operation of a control that can result from unintentional contact.

All errors of the first type are due to confusion of controls because of inadequate distinction among controls. Errors of the second type result from poor design features such as a nonstandard relationship of direction of control movement to expected result, inadequate or ambiguous indications of control position, inadequate feedback from displays, or from a requirement that the control be held in place by the operator, without some signal to that effect. Errors of the third type result from a variety of causes. In almost all cases, they can be eliminated by employing appropriate shields for those controls that can be inadvertently activated.

The probability of each of the above types of errors is largely a function of several PSFs related to the placement and identification of controls, which are listed below.

Performance Shaping Factors

The following PSFs are the most important with respect to errors in the selection or operation of manual controls:

1. Relationship of control to its display (includes physical distance and direction of movement)
2. Identification of control with its function (includes labeling, functional grouping of controls, and use of mimic panels)
3. Specific identification of control (includes control labeling--position, wording, and legibility of label; and control coding--color, shape, size, and position)
4. Anthropometrics (includes ease of reach, ease of visual access, and spacing)
5. Indicators on controls (includes pointers, position marks, and visibility and distinctiveness of indicators)
6. Direction of motion (compliance with populational stereotypes)
7. Operator expectancies (mirror-imaging and symmetry)
8. Immediacy of feedback after control operation
9. Control room layout (includes distance to controls and placement)

Deficiencies Noted

The following deficiencies in layouts of manual controls have been noted at NPPs by the authors and by Seminara et al (1976):

1. There are inadequate means of distinguishing controls: large numbers of identical controls are arranged on panels without any identification other than their labels; controls are not

grouped by function, and no mimic lines are used to help identify the functions of the controls; there is a lack of control coding, such as color or shape, to assist in identifying the controls.

2. Generally, the controls are not arranged in a logical relationship to their displays.
3. In some cases, the distance between a control and its display is such that the operator cannot easily see the display while manipulating the control.
4. In many cases, the controls are difficult to reach and the operator may fail to use them if he is busy.
5. In many cases, the indicators on the controls do not clearly indicate control position.
6. In many cases, there are no provisions for placement of items such as procedures manuals, and the operator will be forced to cover some controls with a manual for want of some place to rest it.
7. Mirror-imaging of control panels presents the same problems in the use of controls as in the use of displays.
8. Labeling of controls is often very terse, and the operator can easily err in control selection.
9. Illumination levels at some control surfaces are barely adequate to read the labels.

The deficiencies noted are such that the operator has difficulty in locating the specific control he requires. In only a few cases does the layout of the control panels help the operator locate and identify the controls. In most cases, the operator has to rely on his knowledge of the control room layout. In many plants, the operators are rotated among control boards for different reactors with different layouts, and in times of stress the operators could easily make errors.

Recovery Factors

In many cases, an error in control activation will be promply corrected by some recovery factor. For example, if an operator turns the wrong control and observes some display that should change, he will soon note that the display has not responded. This usually will alert him to the fact that he did not turn the correct control. In other cases, activation of the wrong control might result in immediate feedback from some other source, such as an annunciator. With certain automatic controls, the system may not accept a change command unless the operator places the control in override. In other instances, the incorrect control may be in the position to which the correct control should be turned, and the status lamp will cue the operator. The HEPs listed below are merely the probabilities that an initial error will be made, without allowance for recovery factors.

Estimated Probabilities of Errors of Commission

This section lists errors of commission only. The probabilities of errors of omission described in Chapter 13, "Valving Operations," apply to controls in general. When a control is being operated in response to an alarm, the error of omission for the control may be disregarded. The benefits of tagging procedures also apply to controls in general. (The use of tags is described in Chapter 15.)

The HEPs listed in Table 12-1 apply to the steady-state operating mode of the plant and to moderately stressful states such as anticipated transients, refueling, and planned shutdown. For conditions associated with high levels of stress, the HEPs should be modified as described in Chapter 17.

The HEPs and error bounds in the table apply to the operation of a single control. In actual operation, only one control is manipulated at a time. In those instances in which controls are handled as pairs, CD is assumed between the two controls, so the HEP for a single control will apply.

Table 12-1. Estimated Probabilities of Errors of Commission in Operating Manual Controls

Task	HEP
Select wrong control in a group of identical controls identified by labels only	.003 (.001 to .01)
Select wrong control from a functionally grouped set of controls	.001 (.0005 to .005)
Select wrong control from a panel with clearly drawn mimic lines	.0005 (.0001 to .001)
Turn control in wrong direction; no violation of populational stereotypes	.0005 (.0001 to .001)
Turn control in wrong direction under normal operating conditions, (design violates a strong populational stereotype)	.05 (.01 to .1)
Turn control in wrong direction under high stress, (design violates a strong populational stereotype)	.5 (.1 to .9)
Set a multiposition selector switch to an incorrect setting. (This error is a function of the clarity with which the indicator position can be determined: the designs of switch knobs and their position indications vary greatly.)	.001 (.0001 to .1)*
Improperly mate a connector. (This includes failures to seat connectors all the way and failure to test the locking features of the connectors for engagement.)	.01 (.005 to .05)

---

*The unusually wide error bounds reflect the wide variety of designs, ranging from good to highly unacceptable ergonomics.

CHAPTER 13. VALVING OPERATIONS

Inasmuch as the proper status of valves is crucial in plant operation and the availability of engineered safety systems, this subject is addressed in detail. We identify two classes of valves: locally-operated valves and motor-operated valves (MOVs). Although both are manually operated, we will use the term manual valve in referring to locally-operated valves and the term "MOV" in referring to valves that are operated by an electrical switch (generally in the control room).

This chapter lists the activities and estimated HEPs associated with (1) changing the valves from their normal operating positions to permit testing, maintenance, calibration, or other work, and (2) subsequent restoration of the valves to their normal operating positions after completion of the work. For convenience, we will use the term maintenance to include testing, calibration, and all other work requiring valves to be changed, change to mean "change the state of a valve from the normal operating position to the nonnormal positon," and restore to mean "restore the state of a valve to the normal operating position."

General Assumptions

The following general assumptions are made regarding plant operating procedures:

1. Operating personnel (including unlicensed reactor operators called "auxiliary operators") generally are responsible for changing the state of a valve. In some instances the operator may valve a system in or out at its boundaries, and maintenance personnel (maintainers) will manipulate valves within the boundaries.

2. The manipulation of valves for maintenance is coordinated between operators and maintainers. Maintainers will not begin work requiring valve changes until the work is cleared by the operators, and operators will not restore valves until the restoration is cleared by the maintainers. Communication between the operators and maintainers may be formal (through a "chain of command") or less formal, depending on individual plant policies.

3. A Level 2 tagging system is assumed. (See Table 15-3, p 15-11 for definitions of tagging systems.) If a Level 1 tagging system is used, for errors of omission we assume CD between all items for which tags are prepared. ZD is assumed with Level 2 and Level 3 tagging systems, except as noted below.

In the case of MOVs, it it standard practice to open the circuit breaker in the Motor Control Center (MCC) after changing the switch at the control panel. The usual practice is to change and tag the control switch, change and tag the circuit breaker, and then tag the valve, and to remove all tags when the valve is restored. With a Level 2 tagging system, if only one valve is to be changed, assume HD between the switch, circuit breaker, and valve. If several valves are to be changed, ZD is assumed within each set of components, (switches, circuit breakers and valves), and also between the sets. ZD is assumed because each set of components is handled as a list in a written procedure, and ZD is usually assumed between steps in written procedures (p 14-9). After the specified switches are changed and tagged, the subsequent steps in the procedures will specify the circuit-breakers to be changed and tagged, and so forth, thus, there will be ZD between all of the steps.

4. Except for major shutdowns and for certain verification tests, it is standard procedure to tag any valve when it is changed.

5. In all cases, it is standard practice to remove all tags when a valve is restored. When a valve has been tagged out for maintenance, there may be several tags on it, one for each maintenance job that has to be done. As each job is completed, only the tag for that job is removed. The valve is not restored until all the jobs have been finished, and all tags removed.

6. Mechanical locks and chains may be used to lock a manual valve in the normal operating position, but never to lock a valve in any other position.

7. Typical NPP procedures are used without requirement for written checkoff of each task as performed. We assume ZD between the steps in the procedures. NPP procedures are discussed in Chapter 14.

8. Typical NPP valve labeling is used, as described on p 3-45.

9. In all cases of valve manipulation, a basic work situation is defined as one in which a single valve is to be manipulated. The valve is in a separate location from the work area, and it is not adjacent to any similar valves. Initially, the HEPs are presented for this simplified situation; then alternative situations are described with their applicable HEPs.

We have omitted very unusual circumstances in the descriptions of alternative work situations. For example, at one of the plants described in WASH-1400, there are eight manual valves on the line that supplies cooling water to the heat exchangers in the Containment Heat Removal System (CHRS). To reach these valves, one must climb down a long, vertical ladder. Plant personnel stated that these valves were NEVER turned.

Despite this assurance, we assigned a 10^{-5} probability that a valve would be inadvertently closed; complete dependence was assumed for the eight valves. Obviously, the estimate is highly subjective, but we must consider the possibility of such an error; and without administratively controlled locks and keys, no recovery factor was allowed for the failure (if it occurred). The above situation is very unusual, and no attempt is made here to account for all the unusual situations that may be discovered.

10. The term recovery factors is used in the limited sense of recovery by human observation, or checking. In most cases we are talking about the probability that one of the people on the job will detect a valving error. We are not addressing the recovery factors that would arise from mechanical or electrical sensors or from the inability of a subsequent task to be performed because of the error. Error probabilities for recovery factors attributable to annunciators and other displays are covered in Chapters 9, 10, and 11. Error probabilities for recovery factors attributable to different tagging systems are included in this chapter.

Specific Tasks

The various functions associated with the manipulations of valves, and the recovery factors attributable to human checking, are listed under the following headings:

Errors of Operators

Errors of Recovery

Errors of Maintainers

Errors of Supervisors

The estimated HEPs for these headings are presented below.

Errors of Operators

Estimated Errors of Commission by an Operator Changing or Restoring Valves

1. Error in writing any item when preparing a list of valves or tags: .003 (.001 to .01)

2. Change or tag wrong valve where the desired valve is one of two or more adjacent, similar-appearing manual valves, and at least one other valve is in the same state as the desired valve, or the valves are MOVs of such type that valve status cannot be determined at the valve itself: .005 (.002 to .02)

3. Restore the wrong manual valve where the desired valve is one of two or more adjacent, similar-appearing valves, and at least two are tagged out for maintenance: .005 (.002 to .02)

4. Reversal Error: It is possible that an operator might attempt to change a valve, switch, or circuit breaker that has already been changed and tagged. This could arise from an error in instructions or from any other cause. In such a case, the operator might fail to notice that the item had already been changed and tagged because of his strong expectancy that the valve state should be changed, and he could make a reversal error restoring the valve instead of changing it.

   a. Reversal Error: Operator "changes" a valve, switch, or circuit breaker that had already been changed and tagged by someone else: .0001 (.00005 to .001)

   b. Reversal Error (as above): If the item had been changed and NOT tagged: .1 (.01 to .5)

5. Failure to note that there is more than one tag on a valve (or circuit breaker or switch) that he has decided to restore: .0001 (.00005 to .0005)

(Ordinarily, an operator would not be instructed to restore a valve until all "holds" had been removed. However, if he were instructed to restore a valve before all "holds" had been removed, there would be an expectancy to find just one tag on the valve.)

6. Change or restore wrong MOV switch or circuit breaker in a group of similar-appearing items: (In case of restoration, at least two items are tagged.) .003 (.001 to .01)

7. Failure to complete a change of state of an MOV of the type that requires the operator to hold the switch until the change is completed, as indicated by an indicator lamp: .003 (.001 to .01)

8. Given that a manual valve sticks, the operator erroneously concludes that the valve is fully open (or closed):

a. Rising-stem valve:

If the valve sticks at about three-fourths or more of its full travel (no position indicator present):

.005 (.002 to .02)

If there is an indicator that shows the full extent of travel: .001 (.0005 to .01)

b. All other valves:

If there is a position indicator on the valve:

.001 (.0005 to .01)

If there is a position indicator elsewhere (and extra effort is required to look at it):

.002 (.001 to .01)

If there is no position indicator:

.01 (.003 to .1)

Estimated Errors of Omission by an Operator Changing or Restoring Valves

1. Omitting any item or tag from a list of valves or set of tags when written procedures or tags are prepared:

.003 (.001 to .01)

2. Failure to carry out a specific oral instruction to change or restore a valve: .001 (.0005 to .005)

For more than one valve, use Pr[F] values for specific oral instructions from Table 14-1, p 14-3.

3. If Level 1 tagging control is used (p 15-11), the set of tags is as compelling as a specific oral instruction. Thus, the estimated HEP to change and tag the first valve in a set of valves is:

.001 (.0005 to .005)

With Level 1 tagging control, for errors of omission there is complete dependence among all items for which tags are prepared -- if the first step in a set is performed, all other steps will be performed.

4. Probability that an operator will omit a particular valve change or restoration when using written procedures: see Table 13-1. If written procedures are not used, see Table 14-1 (p 14-3).

5. Checkoff provision of procedure or checklist is misused; i.e., operator performs several valve operations and then checks off several at a time in the written procedure: .5 (.1 to .9)

6. Plant procedures or policies spelling out details and practices to follow in valve changing will not be followed; shortcuts or personal preferences will be followed instead: .01 (.005 to .05)

Since there is ample history in NPPs of valves left in inappropriate positions having affected the availability of safety systems, the user must evaluate the probability that NPP personnel will not faithfully

Table 13-1. Probability That an Operator Will Omit a Particular Valve in a Sequence when Using Written Procedures

	Short List (< 10 items)	Typical NPP Procedures or Long List (> 10 items)
Checkoff required*	.001 (.0001 to .005)	.003 (.001 to .01)
No checkoff	.003 (.0008 to .01)	.01 (.001 to .05)

*If checkoff is required, the HEPs must be modified by considering the percentage of people who will follow the procedure correctly -- HEP $\geq$ .5 (.1 to .9). If operator performs several operations and then checks off several items at a time, or if he postpones checkoff until job is completed, use HEPs for "No checkoff." Some reviewers (with experience in NPPs or flight crews) of the handbook estimates believe that the probability of misuse of a checklist is considerably higher than .5.

follow plant policies and procedures. A probability of .99 that every person in an NPP intends to carry out plant policies and procedures (whether or not he makes errors of omission or commission in carrying out this intent) would indicate a plant with reasonably good quality control.

7. Failure to tag a valve, circuit breaker, or switch after a change (assuming tags are required): .001 (.0005 to .05)

8. If restored valves are to be locked, this provision can reduce errors of omission, depending upon the level of administrative control of the keys (Table 15-3, p 15-11). Level 2 is assumed to be typical.

    a. Failure to lock a valve after restoration (with Level 2 lock and key control): .003 (.001 to .01)

    b. Failure to lock a valve after restoration (with Level 1 lock and key control): .001 (.0005 to .005)

Errors of Recovery

The only recovery factors addressed in this chapter are those of recovery by human observation or checking. In this section all personnel who are supposed to check the work of someone else are called checkers, whether they are operators or maintainers. When appropriate, a distinction will be made, and the checker will be specified as an operator or maintainer. It is assumed that only operators will check for restoration errors made by other operators.

1. Ordinarily, in checking errors of commission, the HEP for the checker is ten times the HEP of the person performing the task, with an upper limit of .9. (The factor of ten was derived from observations of the differences in error rates between a performer of an activity and a checker of the activity -- McCornack, 1961; and Harris and Chaney, 1967, 1969.)

2. In checking errors of omission, the passive nature of the task and the very high expectancy result in a high HEP:

.1 (.05 to .5)

This does not apply when a checker is specifically told to check a particular item or group of items, in which case the model for compliance with oral instructions applies (Table 14-1, p 14-3).

3. Failure of checker to detect errors of omission or commission on a list of valves or set of tags: .1 (.05 to .5)

Errors of commission in writing items in a written procedure or in filling out tags very often will result in meaningless entries, and will be discovered when the operator tries to use the procedure or tag. No estimate is offered for this recovery factor.

The operation of MOVs often provides immediate feedback in the form of annunciator activation and the changing of indicator lamps. Thus, although the basic HEP of .003 for selection errors with MOVs (p 13-6, item 6) is nearly as high as for manual valves (.005) (p 13-5, items 2 and 3), the automatic and immediate feedback will usually effect prompt recovery of such errors for MOVs. If an annunciator alarms in response to activation of the wrong MOV, the probability that it will be ignored is very small (p 10-9). Also, the status indicator lamps provide prompt feedback of change in status. In those cases in which a legend indicator illuminates, spelling out the designation of the valve that has just been changed, no added recovery factor is allowed for this as a means of detecting that an incorrect valve was selected, as the operator ordinarily would be observing the lamp only to determine that the status had changed.

Reversal errors are less likely to remain undetected with MOVs than is the case with manual valves, because the panel indicators provide unambiguous information about valve position. The same is true in the case of the MOV that is not completely "home" - usually the status lamps will indicate the condition.

4. Failure of checker to detect that a manual valve was not completely "home" after being changed: .5 (.1 to .9)

This high estimate is made because the presence of the tag is the most compelling factor and the checker is likely merely to check that the tag is there.

Allow no recovery factor for a maintainer to detect failure of an operator to fully complete change of an MOV if the status cannot be determined at the valve itself. It is assumed that the maintainer would check that the MOV is tagged and that the circuit breaker is open; he is unlikely to check the status indications in the control room.

5. The estimated HEPs for failure of a checker to note that a valve was not completely "home" after restoration is 10 times that of the operator who made the error, except for valves without position indicators; with which we use a factor of 20, with a maximum HEP of .9. Unlike the case with the comparable valve change error, there is no tag to influence the checker.

In restoring an MOV, if the operator fails to restore a circuit breaker, the error will be detected when he attempts to use the switch on the control panel to restore the MOV. (The joint probability of failing to restore a circuit breaker <u>and</u> of selecting the wrong switch on the control board is judged to be negligible.)

6. Failure of a maintainer to detect a reversal error by an operator changing valves: .5 (.1 to .9)

The estimate for this HEP is high because of the maintainer's very strong expectancy to find the valve changed if there is a tag on it. Even if the operator who made the reversal error followed standard procedures and placed a second tag on the valve, the presence of two tags probably would not alert the maintainer, because multiple tagging is fairly common, and the tag is the most significant cue.

7. If an operator is assigned as checker of a maintainer, use one-half the HEPs of an operator checking another operator. Our rationale is that the dependence will be less between an operator and maintainer than between two operators; i.e., the operator-checker is likely to take more care when checking the work of a maintainer.

8. In some plants, recovery procedures consist of merely checking the paperwork (or counting returned tags). This procedure can catch only some of the possible restoration errors. To be conservative, allow no recovery credit.

Errors of Maintainers

1. Ordinarily, when an operator changes a valve for maintenance, the maintainer will verify the status of the valve before commencing work. Failure of a maintainer to check valve status depends upon a number of factors, the most important of which are:

Personal safety
Nature of instructions
Requirement for written checkoff
Visual accessibility of valve

Personal Safety - If a particular valve or group of valves could endanger the maintainer if it were in the wrong state for maintenance, he would be far more likely to check it than if there were no such safety implications.

Nature of Instructions - Instructions covering the valves to be checked could consist of a short list (10 items or less), a long list (more than 10 items, or a set of typical NPP procedures), or verbal instructions (i.e., no written list).

Checkoff - When using written lists, there may not be a requirement to check off each item as it is verified. The reliability of the human is higher if checkoff is required and done properly (see Table 15-1, p 15-3, and Table 15-2, p 15-9).

Visual Accessibility - If the maintainer is working without a checkoff requirement, he is more likely to verify an item if it is in his field of view. If checkoff is required, visual accessibility of the item is of minor consequence.

Table 13-2 shows the hypothesized interactions of the above factors and the estimated HEPs. The starting point for these HEPs is the maintainer's natural concern for personal safety. If the position of a valve is such that it would endanger him when he worked on the system, our best order-of-magnitude estimate is a probability of .001 (.0005 to .005) that he will fail to check the valve status even in the absence of a written list and even when the valve is not in his field of view. This is similar to the situation of an electrician checking the status of a power switch before working on a circuit -- the same basic HEP of .001 applies. All other HEPs in the matrix are based on rank-orderings of the effects of the listed PSFs, applied to the basic HEP of .001. They are intended

Table 13-2. Probabilities that Maintainer Will Fail to Check Valve Status Before Maintenance*

Type of List Used

Factor	Short List**			NPP Procedures or Long List**			No List+	
Checkoff Used	Yes++	No		Yes++	No			
Within Field of View		Yes	No		Yes	No	Yes	No
Personal Safety Affected								
Yes	.0004	.0006	.0008	.0005	.0006	.0008	.0008	.001
No	.001	.003	.005	.002	.005	.01	.01	.01

NOTES: *In view of speculative nature of the HEPs, estimates of uncertainty bounds are even more speculative. For HEPs greater than 10^{-3}, use a divisor of 5 for the lower bound and a factor of 5 for the upper bound. For HEPs of 10^{-3} or smaller, use a factor of 5 for the upper bound and a constant 10^{-4} for the lower bound.

**Assume ZD between valves.

+For more than one valve, Table 14-1, p 14-3, indicates the HEPs increase as the number of items to remember increases.

++Credit for checkoff must be modified by the percentage of people who fail to use checkoff properly (Table 14-3, p 14-13).

to indicate the direction in which the basic HEP will be modified, rather than the actual extent. Although the entered HEPs are not based on actual data, they are useful approximations, suitable for many applications.

2. In certain cases, after operators have isolated a system by valving it out at the boundaries, it is permissible for maintainers to manipulate valves within the isolated system to accomplish their work. The estimated probability that a maintainer will fail to restore an untagged valve after his work is finished: .05 (.01 to .1) This estimate applies even if he is working with a written list requiring checkoff. Our rationale is that the maintainer's perceived responsibility is maintenance, not restoration. (Untagged valves are stipulated here, because tagged valves are restored by operators.)

Errors of Supervisors

In addition to the errors of operators and maintainers described in the preceding sections, supervisors are subject to errors based on reliance on their staff. Two such hypothetical cases are described below.

1. Supervisor believes that a specific oral instruction he gave an operator to restore a valve was carried out, when it was not:

.90 (.80 to .99)

The high HEP for the supervisor is understandable when considering the very low failure probability of personnel in carrying out a specific oral instruction (Table 14-1, p 14-3). For the case of a single specific oral instruction, the estimated HEP is .001 (.0005 to .005). The expectancy that any single specific oral instruction will be carried out is so great that supervisors seldom check whether the task actually was performed. Usually, the only indication of such a failure would be from a secondary source, which could be a matter of chance. Although we allow no credit for recovery of the supervisor error due to checking, we are allowing a probability of .1 that some secondary source or effect would alert the supervisor that his instruction had not been carried out. Hence, the HEP of .90 (.80 to .99).

In the case of an operator who was instructed to restore a valve: (1) the supervisor just assumed that the operator did what he was told to do, (2) the supervisor asked the operator if he restored the valve, and the operator erroneously said he did (because of a memory error), or (3) some miscommunication occurred between the operator and the supervisor (possibly via a third person) so that the supervisor believed the valve had been restored.

Thus, the probability that a nonrestored valve is thought to be restored can be approximated by the product of .001 (the basic HEP for an operator failing to carry out a specific instruction to change or restore a valve) and .9 (the supervisor's error described above) = .0009. It is known that valves are not always restored when they should be. We usually hear only of those cases in which some serious consequence results, and rarely hear of the cases in which the oversight was caught in time.

2. Another example of a situation in which a supervisor believed that a task was done when it wasn't is the case in which a maintainer was expected to restore a valve upon completion of his work. The estimated error probability for the maintainer is .05 (.01 to .1) (p 13-15, item 2). This estimate is relatively high because the primary concern of the maintainer is maintenance (i.e., fixing something), and restoration of valves is secondary. Maintainers tend to be component-oriented rather than system-oriented. Assuming that a maintainer had not restored a

valve, a supervisor might erroneously believe that it was restored if: (1) the maintainer told his supervisor that he had completed his work and the latter assumed that the valve had been restored, and so informed the shift supervisor, (2) the maintenance supervisor asked the maintainer if he restored the valve, and the latter erroneously said he did (because of a memory error), or (3) some miscommunication occurred among the three people, so that the shift supervisor believed the valve had been restored. In this case the probability that a nonrestored valve is thought to be restored is the product of .05 (the maintainer's HEP) and .9 (the supervisor's HEP) = .045.

Note that in both of the above examples, the influence of the supervisor's probability of recovering an error of the type described is negligible, since normal rounding would restore the combined error probability to the basic error probability (i.e., .0009 rounds to .001 and .045 rounds to .05). Therefore, in cases such as these, the recovery value of the supervisor would be disregarded. Without this kind of conservative approach, an analysis can result in the overoptimistic conclusion that the supervisor will detect all (or at least most) of the critical errors made by his crew.

CHAPTER 14. TASK PROCEDURES

This chapter lists the estimated HEPs and uncertainty bounds associated with oral or written procedures and for arithmetic calculations required for various routine tasks. Errors may occur in the preparation of procedures, failure to use a procedure, or failure to use a procedure as specified. Discussion of the latter two errors overlaps with the subjects of administrative control and recovery factors (Chapter 15).

Oral Procedures

By oral procedures we mean short instructions given by someone in authority to an operator or maintainer. An oral instruction may include one or more items. A typical instruction is, "Restore the valves for System ABC," and is designated as a special instruction item. Here, the recipient must identify the specific valves in that system. Possible problems are: forgetting to initiate the task, looking up the wrong system, finding the correct system in some paperwork but misidentifying one or more valves, and depending on fallible memory rather than writing down the valve designations.

An oral instruction may include more detail than the general instruction above. The instruction may be, "Restore the valves for System ABC; these are valves HC7758, MC1538, and KW8920." In this case, each valve is a special instruction item, and the probability of forgetting to initiate the task is the probability of forgetting to perform one (of only one) special instruction item. Possible errors include those of omission or commission by the person giving the oral instructions, the recipient's forgetting to initiate the task or making errors if he writes down the valve designations, and errors of omission or commission if he relies on memory. Table 14-1 lists estimated HEPs for some of the above errors. Recovery factors will often compensate for an initial error, but such factors are so situation-specific that no general estimates of their effects can be formed.

In the above valve-restoration example, an auxiliary operator may have been told the valve identification numbers for the three valves, but he recalls one of the valve numbers incorrectly by the time he begins the task. Such an error will generally be fully recovered, since it is unlikely that this error will result in an experienced operator's "restoring" the valve with the wrong number. It is likely that the incorrectness of the wrong valve will be obvious. However, if the operator has completely forgotten one of the valves, he will not recover this error of omission. If there is no readout of valve position in the control room, or if the readout is not checked, the error will not be recovered unless it affects other plant functions.

We have not been able to determine the ratio of tasks without written procedures to tasks with written procedures. Brune and Weinstein (1980a) found no mention of a procedure in 24% of almost 1000 Licensee Event Reports (LERs) in the maintenance, test, and calibration areas. In general, these LERs described relatively simple tasks commonly regarded as falling within the skill of the craft, for which detailed written procedures are not provided. For example, a maintainer may be instructed to repair Valve HC1234, and it is assumed that this is all the instruction he needs.

On the basis of their observations at five operating NPPs, Brune and Weinstein concluded that many maintenance tasks are performed without written procedures. A detailed analysis of the LERs indicates that only

Table 14-1. Estimated Probabilities of Errors in Recalling Special Instruction Items Given Orally

Task	HEP
Items Not Written Down by Recipient	
Failure to recall any item, given the following number of items to remember:	
1 (same as failure to initiate task)	.001 (.0005 to .005)
2	.003 (.001 to .01)
3	.01 (.005 to .05)
4	.03 (.01 to .1)
5	.1 (.05 to .5)
Unrecovered failure to recall any item if supervisor asks if that item were performed	Negligible
Items Written Down by Recipient	
Failure to recall each item (exclusive of errors in writing down items)	.001 (.0005 to .005)

NOTE: The above HEP estimates are rounded values taken from Table 8-3, p 8-17.

4% of the tests and 16% of the calibrations were performed without procedures, whereas 37% of all maintenance was performed without procedures -- which suggests that written procedures are less available (or perhaps less useful) for maintenance tasks.

In a substantial number of LERs, unapproved actions were taken even when a written procedure was available. Of 132 maintenance LERs classified as noncompliances, 57 were considered unauthorized. These 57 involved failure to use or follow an available procedure, performing an action not in the procedures, or performing procedural steps out of sequence.

The above data do not tell us how often oral instructions are used, or how often written instructions are ignored. However, they do suggest that there is much reliance on memory, especially during maintenance operations. Table 14-1 indicates that reliance on oral instructions has definite limitations for human reliability.

Written Procedures

HEP estimates related to written procedures are based in part on our observations of several NPPs, on the Brune and Weinstein (1980a) study, and on the Essex study of operator performance at TMI (Malone et al, 1980). The Essex study lists deficiencies typical of procedures in many plants:

- Serious deficiencies in content and format
- Little consistency between nomenclature in procedures and on panel components
- Instructions for control actions that seldom indicate the correct (or incorrect) system response
- Excessive burden placed on operator short-term memory

- Charts and graphs not integrated with the text
- Not clear which procedures apply to which situations
- No formal method for getting operator inputs into updates of procedures
- Grossly deficient instructions for assisting operators in diagnosing the problems related to the TMI accident.

We estimate that the HEPs related to written procedures will be reduced by factors of 3 to 10 if the above types of deficiencies are corrected. The Essex study describes in detail the effects of procedural deficiencies on human errors during the TMI accident.

Preparation of Written Procedures

Three major problems in the preparation of written procedures are the qualifications of the writers, the accuracy of the original procedures, and the provisions for changing existing procedures.

Qualifications of the Writers

Written procedures used in NPPs typically are written by engineers or other highly qualified technical personnel. It is a truism in technical writing that engineers write for engineers, and that their written communications must be "translated" for others. It is not surprising, then, that typical NPP written procedures do not match the capabilities, limitations, and needs of those who use the procedures. Some of these problems are discussed in Chapter 3 (pp 3-47 to 3-52).

Although some utilities hire technical writing organizations to prepare their written procedures, samples of these materials indicate that this practice has not solved the problem of inadequately written materials.

Until some standard set of specifications or guidelines for NPP written procedures can be developed, these problems will continue.

Accuracy of Original Written Instructions

During a simulated LOCA in a dynamic simulator, a highly skilled operator became obviously flustered when he could not locate a critical switch. To cope with the simulated emergency, it is imperative that this switch be operated on a timely basis. After several minutes of running back and forth between sections of the ESF panel, he realized what the problem was. The four-digit identifying number for the switch was incorrect in the procedures. The procedures he was using came from the NPP, not from the simulator facility. As he later noted, people tend to trust the written word, and serious problems can arise when it is incorrect.

Malone, et al (1980, pp 72-74) evaluated the TMI emergency procedure, "Loss of Reactor Coolant/Reactor Coolant System Pressure," in which several problems were found. Several steps critical to handling the accident were not included; e.g., the procedure did not tell the operator what to do if the high-pressure injection system had been initiated automatically. Critical symptoms for leak or rupture, such as a rapid continuing decrease of pressurizer level, were not described. No tolerances were given for critical readings such as the pressurizer level. Of 15 written procedures considered relevant to the accident, only 7 were judged adequate.

There are no means to quantify the probabilities of the above types of inadequacies in written materials. Such errors reflect failure to test the procedures in the actual situation (the simulator exercise described above) as well as failure to anticipate the full scope of situations in which the procedures must be used (the TMI accident). The HEPs below apply to those written procedures that reach the working crew for use on the job, and to errors within the sections that were actually printed. We have no way of determining what additional material should be included, or whether the procedures that are presented are appropriate.

Within these limitations, we estimate a .003 (.001 to .01) probability that an item which is intended to be included in a written procedure will not be included, and the same probability that there will be some error of commission for every item that is included. These figures are our standard estimates for errors of omission and commission; they apply to the preparation of informal lists, maintenance tags, valve restoration lists, etc, as well as to the preparation of formal procedures.

A significant source of error in procedures is the lack of validation of newly prepared or revised procedures. Proper validation requires a trial (or walk-through) of the procedures before release to the users. Often the procedures are released without such validation, and many errors are discovered in use. Required corrections are frequently passed along among the users by word of mouth, and the written procedures themselves remain uncorrected. The error opportunities in such a situation are obvious.

Changes to Written Procedures

Changes to existing procedures constitute a significant source of error. In the Brune and Weinstein (1980a) study of five plants, none of the plants had a systematic process for identifying procedures requiring revision. Reliance was placed upon human memory to identify procedures that might be affected by change (there was no index relating procedures to equipment). Consequently, out-of-date procedures would sometimes be overlooked. We are unable to estimate the probabilities of error in making changes to existing written procedures.

The problem of obsolete procedures is complicated by two factors: the length of time required to update procedures, and the tendency of users to retain "personal" copies of procedures. Some plants mandate revision and review of procedures within 10 days after a change that affects the procedures, whereas other plants do not specify any time requirements for revision and reissue. Obviously, the longer the revision period, the more likely that incorrect procedures will be in use. The other factor in the use of obsolete procedures is the tendency of some personnel to copy procedures which they retain in their personal files and continue to use even after revisions have been issued. As a result, the copy is not accessible for revision. This situation is less likely to happen in plants where personnel must return procedures after use.

One consequence of the above problems is that procedures with handwritten changes will be in use for various periods of time. These handwritten changes are frequently difficult to read, and errors are likely. For example, a numerical value will be crossed out and the new value handwritten between the lines, even in single-spaced copy.

Quality of Written Procedures

This section presents some estimated HEPs based on typical NPP procedures, recommends some improvements to their formats, and presents HEP estimates based on the improved formats.

Quality of Typical NPP Procedures

In Chapter 3, we indicated that typical NPP procedures do not conform to established principles of good writing. The typical format is a narrative style with an excessive number of words to convey essential information -- in engineering parlance, the signal-to-noise ratio is low. Steps in these procedures often include several special-instruction items. The potential problem is that the user may perform one of these special-instruction items, such as, "Check that the pressure is 40 psig ± 5%," look at a gauge to verify this setting, and then return to the written instructions--but to the wrong place and skip an instruction.

Another difficulty with the typical NPP written procedures is that the low signal-to-noise ratio discourages the highly skilled person from using them. It is boring and offensive to read so many words to pick out the few important items. We find that failure to use written instructions properly is often related to the fact that the procedures are poorly written.

Considering only the typical NPP procedures, we have developed the assumptions and estimates described below.

ZD between the steps is assumed, unless an interaction is obvious. For example, if the procedure is written so that Step 5 is to adjust some rotary control, and Step 6 tells what instrument reading should result from that adjustment, the assumption of ZD between the two steps is clearly inappropriate. Apart from such obvious cases, ZD is usually assumed among the steps in a written procedure.

The estimates in Table 14-2 are made on the assumption that the procedures are being used correctly. For modifications for incorrect use, see Table 14-3, p 14-12.

Improved Written Procedures

Several possible formats could be employed for NPP procedures. We favor the columnar format shown in Figure 3-12 (p 3-51). The signal-to-noise ratio is high, each numbered item has only one special instruction item (or completely dependent set of items -- such as the four switches in Step 2 in the figure), and there is provision for checking off each step as it is performed. The "Check" column is placed next to the "Step" column so that the user (or later checker) can quickly run down the two columns and see which steps might have been omitted. Thus, while placement of the "Check" column at the right side of the sheet seems logical in view of our normal left-to-right reading, it is more convenient to place this column as shown in Figure 3-12.*

Extrapolating from the Haney study (1969) described on p 3-50, we estimate a factor of 3 reduction in errors of omission and commission if the columnar format shown in Figure 3-12 is used instead of the typical narrative style of NPP procedures. There should also be some gain in the number of people who use the procedures properly.

---

*As a training aid, we recommend combining the columnar format with any necessary background or explanatory detail in a narrative format, so that a trainee can read the text, perform the required task, and check it off -- thus acquiring the habit of using the procedures correctly as part of the training program.

Table 14-2. Estimated Probabilities of Error When Using Written Procedures Correctly

Task	HEP
NonPassive Tasks Such as Maintenance, Test, or Calibration	
Errors of Commission	.003 (.001 to .01)
Errors of Omission	
- Procedures with no checkoff provisions (assuming correct use of procedures)	
- Short list ($\leq$ 10 special instruction items)	.003 (.001 to .01)
- Long List (> 10 special instruction items)	.01 (.005 to .05)
- Procedures with checkoff provisions (assuming correct use of checkoff)	
- Short list ($\leq$ 10 special instruction items)	.001 (.0005 to .005)
- Long list (> 10 special instruction items)	.003 (.001 to .01)
Passive Tasks Such as the Walk-Around Inspection	
Failure to recognize an incorrect status when checking each item as he looks at it	.01 (.005 to .05)
Failure to recognize an incorrect status when checking off several items after looking at several	.1 (.05 to .5)
Failure to recognize an incorrect status if no checking provision is required	Use Walk-Around HEPs in Chapter 8

Table 14-3. Estimated Probabilities of Nonuse and Misuse of Written Procedures

Task	HEP
Nonpassive Tasks Such as Maintenance, Test, Calibration, or Control Room Tasks	
Nonuse of Written Procedures	
- Maintenance	.3 (.05 to .9)
- Valve change or restoration	.01 (.005 to .05)
- Test or calibration	.05 (.01 to .1)
- Control room	.05 (.01 to .1)
Misuse of Written Procedures	
- Checkoff required but misused (checking off several items after looking at several)	.5 (.1 to .9)
- Checkoff required but not used, or checking off all items after job is finished (Note: Consider this action equivalent to *no* provision for checkoff)	.01 (.005 to .05)
Passive Tasks Such as the Walk-Around Inspection	
Procedures available but not used	.05 (.01 to .1)
Checkoff required but misused (checking off several items after looking at several)	.5 (.1 to .9)*
Checkoff required but not used, or checking off all items after job is finished (Note: Consider this action equivalent to *no* provision for checkoff)	.01 (.005 to .05)

---

*Some reviewers of the handbook have stated that this assumption of correct use of a checklist 50% of the time is overly optimistic.

Even if the written procedures use a narrative style, there is ample room for improvement. Brune and Weinstein (1980b) have developed a checklist for use in evaluating a plant's written procedures. Their checklist and instructions for its use are appended to this chapter. We estimate that if a written procedure conforms to the checklist criteria, probabilities of errors of omission and commission will be reduced by a factor of 3. We further estimate that, if the columnar format is incorporated along with the checklist criteria, there will be a further reduction of relevant HEPs.

Use of Available Written Procedures

A common error in a human reliability analysis is to assume that available written procedures will be used and used properly. Even in work with severe penalties for nonuse or incorrect use of written procedures, nonuse and misuse have been observed. In the absence of such penalties, it is reasonable to assume a greater frequency of such practices. At every plant Brune and Weinstein visited, at least one manager or supervisor expressed concern that personnel might not use or follow written procedures as intended. The nonuse or misuse of written procedures is obviously related to the quality of administrative control in a plant (discussed in the next chapter). However, these problems are also related to the quality of the written materials themselves.

Considering the usual poor quality of written materials in NPPs, we have developed the estimated HEPs in Table 14-3. (There are no objective data to back up these estimates; they are based on our experience.) As can be seen from the HEPs in the table, we are not optimistic about the use of procedures by maintenance personnel.

Arithmetic Calculations

Planners of work operations often overestimate the accuracy with which technical people may perform routine arithmetic calculations. The relatively large HEP for such calculations is illustrated by a study in which experienced inspectors measured the locations of holes in a calibrated steel test plate (Rigby and Edelman, 1968b). Using micrometer-type instruments, 12 inspectors measured X and Y coordinates for each of 9 holes, for a total of 216 6-digit data points. Typically, each inspector made more than one measurement for each value, so that there were different numbers of opportunities for each inspector to make an error. When taking measurements, the inspector had to read values from a gauge and a meter and assign a plus or minus sign as appropriate. As he read each value, he wrote it on a work sheet, and then performed additions, subtractions, and divisions to arrive at a 6-digit value for an X or Y coordinate. In performing this simple arithmetic the inspectors made a total of 9 errors in 698 opportunities, an HEP of about .01.

In earlier studies, HEPs per digit of about .002 for addition, subtraction, and division, and about .003 for multiplication were obtained for university students acting as experimental subjects (Weldon, 1956; and Weldon et al, 1955). In another study of computational accuracy, the performance of naval cadets resulted in HEPs per digit of about .007 for addition and .016 for subtraction (Trumbull, 1952a and b). Obviously, the experimental conditions and subjects are different for the three studies. In the studies using students and naval cadets, speed was emphasized along with accuracy, and their only task was the set of required computations. In the inspector study, accuracy was emphasized, and the calculations were only one part of the task performed. In this handbook, we use an overall HEP of .01 (.005 to .05) for arithmetic computations in general. This estimate applies to the simple arithmetic calculations that would be required for calibration and maintenance tasks.

APPENDIX TO CHAPTER 14. PROCEDURES EVALUATION CHECKLIST FOR MAINTENANCE, TEST, AND CALIBRATION PROCEDURES

The following checklist for maintenance, test, and calibration procedures and the instructions for use of the checklist are from Brune and Weinstein (1980b). Most of the items on the checklist apply to all kinds of NPP procedures, not just to maintenance, test, and calibration. However, the above authors are currently developing a checklist for evaluating emergency procedures. This latter checklist will be available around March 1981.

The procedures evaluation checklist, developed by Human Performance Technologies, Inc. as part of an NRC/Sandia human factors study of NPP operations and procedures, can be used to identify deficiencies in maintenance, test, and calibration procedures that can lead to errors in performance. The checklist was developed on the basis of the authors' experience, detailed observation, interviews with personnel at five NPPs, and an analysis of the 751 abstracts of LERs submitted by all plants during the 4-year period 1975-1978 that were attributed to procedural deficiencies.

The checklist can:

(1) Identify deficiencies in a set of procedures with the objective of correcting them, and/or

(2) Identify deficiencies in a set of procedures with the objective of correcting the process that produced the deficient procedures.

Following the checklist is a set of instructions for its application. The evaluator of a <u>licensee's</u> written procedures must be familiar with the procedures and operations of the plant to which the checklist is to be applied.

PROCEDURES EVALUATION CHECKLIST

(Maintenance, Test, and Calibration Procedures)

Procedure Title/No.____

Revision__________Reviewed by______________________Date______

Review the procedure for each of the following characteristics. If it possesses the characteristic, check <u>Yes</u>; if it lacks the characteristic, check <u>No</u>. Check N/A (Not Applicable) if the characteristic does not apply to the procedure.

The ratings (A,B,C,D) indicate the relative impact of the characteristic on the performance of personnel using the procedures. If a procedure lacks a characteristic rated <u>A</u>, performance error is most likely; if it lacks a characteristic rated <u>D</u>, performance error is least likely.

Evaluation Method: Perform a document review of procedure and referenced documents.

Item	Rating	Yes	No	N/A
1. Does each page provide the following identification information?				
1.1 Procedure number and/or title	D	☐	☐	☐
1.2 Date of issue	D	☐	☐	☐
1.3 Revision number	D	☐	☐	☐
1.4 Page number	D	☐	☐	☐
2. If this is a temporary procedure, is it clearly marked with the expiration date?	D	☐	☐	☐
3. Is the last page of the procedure clearly identifiable by marking; e.g., Page____ of____, Final Page?	D	☐	☐	☐
4. Does the procedure have a unique and permanently assigned number? That is, if the procedure becomes deleted will the number be retired rather than reassigned?	D	☐	☐	☐

Item	Rating	Yes	No	N/A
5. Does the procedure provide a statement of purpose or brief description which clearly specifies the function it performs in an introductory section preceding the instructions?	D	☐	☐	☐
6. Does the procedure provide the following job planning information in an introductory section preceding the instructions?				
6.1 Other actions or procedures which must be completed prior to use	B	☐	☐	☐
6.2 Plant, system or equipment conditions which must exist prior to use	B	☐	☐	☐
6.3 Precautions which must be observed in the performance of the procedure	C	☐	☐	☐
6.4 The specific equipment (by part number and/or unique nomenclature) to which the procedure is applicable	D	☐	☐	☐
6.5 Special tools and test equipment required to perform procedure (by part number and/or unique nomenclature)	D	☐	☐	☐
6.6 Other documents; e.g., procedures, drawings, schematics, required to perform procedure	C	☐	☐	☐
6.7 Qualifications of personnel permitted to perform procedure	C	☐	☐	☐
6.8 Number of personnel required to perform procedure	D	☐	☐	☐
7. Are the titles and numbers of all referenced documents identified correctly?	C	☐	☐	☐
8. Are all of the documents referred to in the instructions listed in the introductory section of the procedure?	D	☐	☐	☐
9. Does one procedure provide all of the instructional information necessary to perform the activity (rather than refer personnel to other procedures to perform parts of the activity)?	C	☐	☐	☐
10. If the procedure refers personnel to other procedures for instructional information, does it specify the applicable sections, paragraphs, or pages?	D	☐	☐	☐
11. If the procedure refers personnel to other procedures for instructional information, each reference must be evaluated as an independent procedure starting with Step 1 of the checklist.	-			
12. Does the procedure provide adequate quality control (QC) hold points?	C	☐	☐	☐

Item	Rating	Yes	No	N/A
13. Does the procedure provide for verification and signoff of actions?	B	☐	☐	☐
14. If yes, is every step signed off or initialled?	D	☐	☐	☐
15. If yes, are the verifications usually performed by persons other than those performing the action?	C	☐	☐	☐
16. If the procedure refers to a skill-of-the-craft task; i.e., is general rather than specific, go directly to Step 34 to evaluate.				
17. Evaluate the complexity of the action instructions by determining the average number of actions (verbs) called out per step. Base estimate on a sample of 20% of the steps, or, if the sample size is less than 10, use all steps. Is the average number of actions per step 1.5 or less?	C	☐	☐	☐
18. Are approximately 90% or more of the instructions written in short, concise, identifiable steps (as opposed to multi-step paragraphs)?	C	☐	☐	☐
19. Evaluate the level of specificity of a procedure by determining the percent of steps in a selected sample that meet <u>all</u> of the following criteria. 19.1 The action to be taken is specifically identified (open, close, torque, etc) 19.2 Limits (if applicable) are expressed quantitatively (2 turns, 100 inch-lbs., etc) 19.3 The equipment or parts are identified completely (HPCI-MO-17, etc) Base the estimate on a sample of 20% of the steps in a given procedure or a minimum of 10 steps. Do at least 90% of the steps evaluated meet all the above criteria?	B	☐	☐	☐
20. If precautions or explanations are applicable to the performance of specific steps or series of steps, are they placed immediately ahead of the step(s) to which they apply?	C	☐	☐	☐
21. If more than one person is required to perform the procedure, is the procedure written to one 'primary' user; that is, is it clear from the way that instructions are written that one person is responsible for coordinating the activity?	C	☐	☐	☐
22. Are graphs, charts, and tables adequate for readability and interpolation or extraction of values?	C	☐	☐	☐

Item	Rating	Yes	No	N/A
23. Are worksheets designed to facilitate all required computations? That is, are spaces provided for recording all data and processing them (performing additions, multiplications, etc)?	C	☐	☐	☐
24. Does the procedure (or related data sheets or worksheets) provide for the independent verification and signoff of computations?	C	☐	☐	☐
25. If qualitative acceptance criteria are used, should they be restated in quantitative terms?	C	☐	☐	☐
26. If quantitative acceptance criteria are used, are they stated as a range with a midpoint as opposed to a point value?	C	☐	☐	☐
27. Are the acceptance criteria compatible with the limits expressed in requirements documents?	A	☐	☐	☐
28. If computations are required by the procedure, are they based on technically accurate, complete, and up-to-date formulae?	A	☐	☐	☐
29. If items (valves, breakers, relays, solenoids, jumpers, fuses, switches) require alignment to perform the procedure, do the alignment instructions in the procedure meet <u>all</u> of the following criteria?	B	☐	☐	☐
29.1 Each item requiring alignment is individually specified. (Note--It is not acceptable to refer personnel to previous steps)				
29.2 Each item is identified with a unique number or nomenclature				
29.3 The position in which the item is to be placed is specified				
29.4 The position in which the item is placed is verified and signed off				
30. If any of the above alignment instructions are for system restoration, do they meet <u>both</u> of the following verification criteria?	B	☐	☐	☐
30.1 The position of each item is verified by signoff				
30.2 The verification is performed by someone other than the person performing the alignment				
31. If any follow-on action, test, or procedure must be performed upon the completion of this procedure, does the procedure or a related document (e.g., work order) instruct the user regarding what follow-on action is required and whom to notify?	C	☐	☐	☐

Item	Rating	Yes	No	N/A
32. Does the procedure provide instructions for reasonable contingencies? For example, if equipment is operating outside the range specified by the procedure, is the person instructed what action to take?	C	☐	☐	☐
Evaluation Method: Perform a walk-through of procedure.				
33. Are equipment numbers and/or nomenclature used in the procedure identical to those which are displayed on the equipment?	B	☐	☐	☐
Evaluation Method: Observe licensee representative perform a walk-through of procedure.				
34. Determine whether the amount and kind of information (level of detail) provided by the procedure is adequate for the intended users. Make this evaluation by observing a user representative of the least qualified personnel permitted to perform the procedure simulate a walk-through of the procedure. Are the following criteria met?				
34.1 Can the procedure be performed in the sequence in which it is written?	A	☐	☐	☐
34.2 Can the user locate and identify all equipment referred to in the instructions?	A	☐	☐	☐
34.3 Where general rather than specific instructions are provided, can the user explain in detail how to perform the general instruction?	A	☐	☐	☐
34.4 Can the user perform the procedure without obtaining additional information from persons or documents not specified by the procedure?	B	☐	☐	☐
34.5 Can the user perform the procedure without obtaining direct assistance from persons not specified by the procedure?	B	☐	☐	☐

Action: ____

____

____

____

Item	Rating Yes No N/A

Disposition: ____

____

____

____

DISCUSSION OF CHECKLIST

Item Format

The procedures evaluation criteria are expressed in question form so that they can be answered by Yes or No. They are constructed so that a Yes answer indicates that the procedure possesses a desirable characteristic. A No answer indicates a procedural deficiency. In some cases, it will not be possible to evaluate a procedure on a characteristic because it is not applicable to the procedure. For example, a procedure cannot be evaluated regarding the accuracy with which it identifies reference documents if it does not refer to other documents. In this case, check Not Applicable rather than leave the item unanswered.

Explanation of Checklist Items

The checklist items are listed below. In cases in which the relationship between a procedural characteristic and the quality of human performance might not be apparent, an explanation is provided.

1. Does each page provide the following identification information?

Item Rating Yes No N/A

1.1 Procedure number and/or title

1.2 Date of issue

1.3 Revision number

1.4 Page number

Explanation None

2. If this is a temporary procedure, is it clearly marked with the expiration date?

Explanation None

3. Is the last page of the procedure clearly identifiable by marking; e.g., Page____of____, Final Page?

Explanation The last page of a procedure is most vulnerable to becoming detached and lost. It should be made obvious to the user if the last page is missing.

4. Does the procedure have a unique and permanently assigned number? That is, if the procedure becomes deleted will the number be retired rather than reassigned?

Explanation When a procedure becomes obsolete, some licensees reissue its number to a new procedure that may have entirely different subject matter. This practice can introduce error into their cross-referencing system as well as make it difficult for an auditor to retrace an event.

5. Does the procedure provide a statement of purpose or brief description which clearly specifies the function it performs in an introductory section preceding the instructions?

Explanation None

6. Does the procedure provide the following job planning information in an introductory section preceding the instructions?

   6.1 Other actions or procedures which must be completed prior to use

   6.2 Plant, system, or equipment conditions which must exist prior to use

   6.3 Precautions which must be observed in the performance of the procedure

   6.4 The specific equipment (by part number and/or unique nomenclature) to which the procedure is applicable

Item Rating Yes No N/A

6.5 Special tools and test equipment required to perform the procedure (by part number and/or unique nomenclature)

6.6 Other documents; e.g., procedures, drawings, schematics, required to perform procedure

6.7 Qualifications of personnel permitted to perform procedure

6.8 Number of personnel required to perform procedure

Explanation None

7. Are the titles and numbers of all referenced documents identified correctly?

Explanation None

8. Are all of the documents referred to in the instructions listed in the introductory section of the procedure?

Explanation None

9. Does one procedure provide all of the instructional information necessary to perform the activity (rather than refer personnel to other procedures to perform parts of the activity)?

Explanation Several studies have shown that referencing to other documents is a major source of complaint. In some cases, they may not be obtained and used.

10. If the procedure refers personnel to other procedures for instructional information, does it specify the applicable sections, paragraphs or pages?

Explanation None

11. If the procedure refers personnel to other procedures for instructional information, each reference must be evaluated as an independent procedure starting with Step 1 of the checklist.

Explanation All documents that constitute the procedural information system should meet the same human factors criteria to reduce performance errors. In many cases, the documents vary greatly with respect to format and quality of content.

12. Does the procedure provide adequate QC hold points?

Explanation None

Item Rating Yes No N/A

13. Does the procedure provide for verification and signoff of actions?

Explanation See item 15

14. If yes, is every step signed off or initialled?

Explanation See item 15

15. If yes, are the verifications usually predominantly performed by persons other than those performing the action?

Explanation Verification is the primary method for ensuring compliance with procedures. Self-verification by checking, initialling or signing steps serves as an aid or reminder to the procedure user to perform the step. However, it is too easily subject to abuse to serve as a compliance control. If it is important to ensure compliance with an action because of the consequences of a performance error, verification by someone other than the person performing the action is in order. It is required if an error would otherwise remain undetected.

16. If the procedure refers to a skill-of-the-craft task; i.e., is general rather than specific, go directly to Step 34 to evaluate.

Explanation Many procedures that are intended for use by skilled craftsmen lack detail. The performance of Step 34 enables the procedure evaluator to check the adequacy of a procedure with respect to the qualifications of the personnel permitted to use it.

17. Evaluate the complexity of the action instructions by determining the average number of actions (verbs) called out per step. Base estimate on a sample of 20% of the steps, or, if the sample size is less than 10, use all steps.

Is the average number of actions per step 1.5 or less?

Explanation The complexity index (CI) of a procedure is defined as the average number of actions stated in the instructional steps or paragraphs. The average is computed from a random sample of steps or paragraphs in a procedure.

$$CI = \frac{\text{Number of Actions in a Sample of Steps or Paragraphs}}{\text{Number of Steps or Paragraphs Sampled}}$$

The number of actions is simply the number of verbs in a step or paragraph. For example, the instruction

Item

Rating Yes No N/A

"Turn switch XXX to position No. 2, observe value on pressure gauge XX and record value" has three actions. The more actions that are expressed, the less likely they will be recalled accurately, particularly if they are unrelated actions. Ideally, a step should contain only one action unless the actions are related, in which case up to three actions in a step are acceptable. Related actions are a group of actions required to produce a single result. The example illustrates related actions. Their single object is to obtain a value.

18. Are approximately 90% or more of the instructions written in short, concise, identifiable steps (as opposed to multi-step paragraphs)?

    Explanation None

19. Evaluate the level of specificity of a procedure by determining the percent of steps in a selected sample that meet all of the following criteria.

    19.1 The action to be taken is specifically identified (open, close, torque, etc)

    19.2 Limits (if applicable) are expressed quantitatively (2 turns, 100 inch-lbs., etc)

    19.3 The equipment or parts are identified completely (HPCI-MO-17, etc)

Base the estimate on the sample of 20% of the steps in a given procedure or a minimum of 10 steps.

Do at least 90% of the steps evaluated meet all the above criteria?

Explanation The above criteria list the basic characteristics of a specific (versus general) instruction. Fewer errors of interpretation or omission result from instructions with high specificity.

20. If precautions or explanations are applicable to the performance of specific steps or series of steps, are they placed immediately ahead of the step(s) to which they apply?

    Explanation None

21. If more than one person is required to perform the procedure, is the procedure written to one 'primary' user, that is, is it clear from the way that instructions are written that one person is responsible for coordinating the activity?

    Explanation Almost all instructions examined by the study team did not identify the number of

Item Rating Yes No N/A

personnel involved in performing an activity. In many procedures it was possible to infer that the coordinated efforts of two or more persons were required. However, it was often difficult or impossible to determine which party was being instructed by a step. To the extent that the activity is not structured, errors in communication and omissions of actions can result.

22. Are graphs, charts, and tables adequate for readability and interpolation or extraction of values?

Explanation Misinterpretation of graphs, charts, and tables has resulted in performance errors. It is often traceable to poor readability of these materials--which, in turn, is attributable to 1) inadequate reproduction or 2) inadequate original construction. The following guidelines are provided to evaluate readability.

Reproduction--In some cases, copies are so many generations removed from the original or master copy that lines in graphs, charts, and tables have deteriorated or disappeared, making it difficult to track or interpolate values. Letters and numbers can undergo similar deterioration. Also, materials have sometimes been reduced in size so that readability is impaired. Letters and numbers should be at least 1/8 in. in height, unbroken and unfilled. All lines in the reproductions should be as visible as they are in the original or master copies. First, compare the reproductions to the originals or master copies. Then evaluate the readability of the reproductions under the conditions of illumination in which personnel use them.

Original construction--Letters and numbers should be typed rather than handwritten. Lines on graph paper should be reproducible on licensee reproduction equipment. On graphs, units of measurement used in plotted values should be compatible with divisions on graph paper. That is, if plotted values progress in units of five; e.g., 5, 10, 15, etc, it is better to separate the values by five lines than by four lines. To facilitate accuracy of locating values in charts and tables look for such aids as 1) partitioning tables with lines, 2) arranging values in subgroups; e.g., inserting spaces between subgroups of five values, and 3) placing connecting lines between values or nomenclature and values.

Item Rating Yes No N/A

23. Are worksheets designed to facilitate all required computations? That is, are spaces provided for recording all data and processing them (performing additions, multiplications, etc)?

Explanation None

24. Does the procedure (or related data sheets or worksheets) provide for the independent verification and signoff of computations?

Explanation None

25. If qualitative acceptance criteria are used should they be restated in quantitative terms?

Explanation None

26. If quantitative acceptance criteria are used are they stated as a range with a midpoint as opposed to a point value?

Explanation When equipment does not permit the setting of point values, or when a range of values is acceptable, the acceptance criteria should be expressed in terms of ranges. However, they should be expressed in a form to avoid errors of addition, subtraction or conversion. Example:

Preferable

____________to____________

Best

________ (____________ ________________)
midpoint lower limit upper limit

Not Preferable

____________±____________

Worst

____________±_____%

27. Are the acceptance criteria compatible with the limits expressed in requirements documents?

Explanation None

28. If computations are required by the procedure, are they based on technically accurate, complete, and up-to-date formulae?

Explanation None

29. If items (valves, breakers, relays, solenoids, jumpers, fuses, switches) require alignment to perform the procedure, do the alignment instructions in the procedure meet all of the following criteria?

Item Rating Yes No N/A

29.1 Each item requiring alignment is individually specified. (Note--It is not acceptable to refer personnel to previous steps)

29.2 Each item is identified with a unique number or nomenclature

29.3 The position in which the item is to be placed is specified

29.4 The position in which the item is placed is verified and signed off

Explanation Two of the primary contributors to misalignment are lack of specificity of instructions and lack of physical verification of position. The criteria listed above are aimed at improving specificity and verification. In some procedures it was found that instructions were adequate for initial alignment but shortchanged realignment by simply directing personnel to "Reposition valves listed in Step 5." In this instance, personnel were not provided a means within the procedure for verifying valve positions. The instruction should have relisted the valves, their new positions and provided signoff for each valve.

30. If any of the above alignment instructions are for system restoration, do they meet both of the following verification criteria?

30.1 The position of each item is verified by signoff

30.2 The verification is performed by someone other than the person performing the alignment

Explanation It was found that up to three-fourths of undetected alignment errors occur during restoration. Independent physical verification of position is less likely to be performed during this process. The requirement for independent verification is aimed at reducing this error. The independent verification should involve physically checking the positions--not be confined to simply checking log entries and tags.

31. If any follow-on action, test, or procedure must be performed upon the completion of the procedure, does the procedure or a related document (e.g., work order) instruct the user regarding what follow-on action is required and whom to notify?

Explanation None

Item Rating Yes No N/A

32. Does the procedure provide instructions for reasonable contingencies? For example, if equipment is operating outside the range specified by the procedure, is the person instructed what action to take?

Explanation Many procedures are written as though all acceptance criteria will be met. They do not address the exceptions. Personnel should be instructed within the procedure what actions to take in the event criteria are not met.

33. Are equipment numbers and/or nomenclature used in the procedure identical to those which are displayed on the equipment?

Explanation None

34. Determine whether the amount and kind of information (level of detail) provided by the procedure is adequate for the intended users. Make this evaluation by observing a user representative of the least qualified personnel permitted to perform the procedure simulate a walk-through of the procedure. Are the following criteria met:

34.1 Can the procedure be performed in the sequence in which it is written?

34.2 Can the user locate and identify all equipment referred to in the instructions?

34.3 Where general rather than specific instructions are provided, can the user explain in detail how to perform the general instruction?

34.4 Can the user perform the procedure without obtaining additional information from persons or documents not specified by the procedure?

34.5 Can the user perform the procedure without obtaining direct assistance from persons not specified by the procedure?

Explanation It is important to evaluate whether or not procedures are adequate for use by qualified personnel. Because there are inadequate definitions of adequate procedures or of personnel qualifications, this assessment cannot be made definitively. There is considerable room for different interpretations and disagreement between inspectors and licensees. The above test permits an objective assessment of procedural adequacy.

Item Ratings

The ratings A, B, C, or D indicate the impact of an item on the quality of human performance. If a procedure is deficient with respect to the characteristic referred to by the item, a performance deviation is more likely to occur than if the procedure possesses the characteristic. The absence of some procedural characteristics is more likely to result in performance deviations than the absence of others. It is therefore necessary to develop a method of rating the checklist items to indicate to the evaluator the relative importance of the characteristic stated in the item. The rating considerations are shown in Table A14-1. These ratings, integrated with the evaluator's own knowledge of the consequences of error associated with the performance of a specific procedure or action, should enable him to assess the importance of correcting a particular procedural deficiency. In general, it should be considered mandatory to correct a deficiency rated A or B. Correction of a deficiency rated C may or may not be considered mandatory, depending upon the evaluator's judgment regarding the consequences of error and situational stress factors associated with use of the procedure. A rating of D would not ordinarily be regarded as a mandatory change. However, correction is desirable if the intent is to reduce the

Table A14-1. Rating Scale for Procedural Deficiencies

	Probability of Performance Deviation Under:		
Rating	Low Stress	Moderate Stress	High Stress
A	Moderate	High	High
B	Moderate	Moderate	High
C	Low	Moderate	Moderate
D	Low	Low	Moderate

Description of Ratings

A-Errors are likely to occur during low stress (normal) conditions and will be made frequently under moderate and high stress conditions.

B-Errors are likely to occur during low and moderate stress conditions and will occur frequently under high stress.

C-Errors are not very likely under low stress but could readily occur under moderate and high stress.

D-Errors are not very likely to occur under low and moderate stress but could readily occur during high stress.

frequency of performance error to the minimum attainable by means of procedures.

Evaluation Methods

The checklist employs three methods for evaluating procedures. The Document Review method is used to evaluate a procedure on Items #1 through #32. The Walk-Through Method is used for a procedure on Item #33 and the Observation Method applies to Item #34. The methods are described below.

Document Review This method consists of collecting a sample of the procedures of interest and their related documents and then examining their contents and interrelationships. Typically, related documents will consist of 1) all drawings, procedures, schematics, etc specifically referred to by the procedure, 2) technical specifications, Final Safety Analysis Reports (FSARs), and other basic requirements documents which reasonably might affect the content of the procedures, and 3) corporation policies and station directives dealing with procedures contents, development, and implementation. These documents together comprise the information system affecting the performance of a maintenance, test, or calibration activity. If an evaluator cannot evaluate a characteristic from the available documents alone when a document review has been specified, it can be assumed that the information system is deficient with respect to completeness or with respect to organization. Either deficiency will affect the quality of procedural content adversely. At the least, an information system must be auditable.

Walk-Through Some evaluations, such as determining the correspondence between equipment nomenclature or identification numbers used in a procedure and the nomenclature or numbers actually displayed on equipment, can be performed only by walking through the facility with the procedure in hand and comparing the two. During the walk-through it might be desired to make selected human factors observations of the work environment, the facility layout, and the equipment, all of which bear upon the effectiveness and safety of personnel performance. For example, the evaluator might wish to assess the readability of legends and displays from the perspective of the person performing the procedure.

Observation Unlike the preceding methods, the performance of this evaluation requires the direct support of licensee personnel. The objective of this method is to judge whether the amount and kind of information provided by the procedure is complete with respect to the information needs of the user. That is, the evaluator seeks to evaluate the adequacy of the "level of detail" of the procedure. This attribute of a procedure is the most difficult of all procedural characteristics to evaluate. Judgments of adequacy of level of detail are based on assumptions about the qualifications of the personnel for whom the procedure is provided. Such assumptions are often tenuous at best because, unlike operators, documentation detailing the qualifications of personnel who perform these procedures--particularly maintenance procedures--is inadequate or nonexistent.

To reduce the probability of human error in the performance of an activity, a procedure must be designed to be usable for the least qualified person permitted to use the procedure. This requirement implies that the procedure must provide all of the information needed by persons representative of that skill level and, furthermore, must express the information in understandable language (vocabulary, sentence structure). Partial evidence of the completeness of a procedure can be obtained by observing a person who is representative of the minimum skill level perform a walk-through of the procedure, simulating the actions specified in the instructions.

Typically, a procedure is composed of general and specific instructions. The user should be able to explain in detail how to perform a general instruction. The user should be able to perform the entire procedure without seeking information from other personnel, unless they are specified by the procedures, and without referring to documents that are not specified by the procedures. If either of these criteria is not met, the procedure is incomplete.

CHAPTER 15. RECOVERY FACTORS AND ADMINISTRATIVE CONTROL

There are many recovery factors that can prevent human error from producing unwanted effects in NPPs. Some of these are 100% effective. For example, in some cases Step 5 cannot be performed unless Step 4 is done; in such a case, omission of Step 4 is always recovered when Step 5 is attempted. In risk assessment we are not too concerned with such recovery factors unless time constraints are very tight.

Most recovery factors in NPPs are based on personnel interaction or on information fed back to the operating personnel via displays. This chapter primarily addresses recovery factors afforded by human interaction. We use the term human redundancy to refer to various levels of self-checking as well as to the use of a second person to check performance. Different recovery probabilities are estimated for different levels of human redundancy. The experimental literature and related studies behind these estimates are described.

This chapter also addresses the very important recovery factor of administrative control. This term refers to the kinds of checking of human performance mandated in a plant and the extent to which plant policies are carried out and monitored, including tagging controls and lock and key controls for valves and other components. Thus, administrative control affects the effectiveness of recovery factors as well as the likelihood of initial human errors.

Human Redundancy

There is no human redundancy if one person performs some task and no one (not even the person himself) checks his work. Usually more than one person is assigned to tasks such as maintenance, calibration, and fuel exchange. The probability of error recovery is determined by the extent to which there are opportunities to detect an error, the extent to which those opportunities are exercised, and the accuracy of those who check the work that has been performed.

We take the conservative view that not much, if any, recovery credit should be given for a person checking his own work. Such cases should be handled individually and some assessment made of the level of dependence between a person's errors and the probability that he will catch these errors (see Chapter 7). This chapter deals only with cases in which one person checks another's work.

Table 15-1 presents estimated HEPs for a person who inspects or monitors another's work in an NPP. The terms in the table are defined below. In all cases, we assume that the state of affairs to be monitored is directly observable -- no interpretation is required. The probability that the monitoring behavior will indeed be carried out is discussed under the topic of "Administrative Control" later in this chapter. (The terms checker, <u>inspector</u>, and <u>monitor</u> all refer to someone who checks another's work, either while that person is doing the work or after its completion.)

With no objective data on the HEPs of checkers in NPPs, we base the estimates in Table 15-1 on (1) extrapolations from a series of experiments and studies of inspectors in industrial processes (cf Harris and Chaney, 1967 and 1969; McCornack, 1961; McKenzie, 1958; and Rigby and Swain, 1975) and (2) our experience in weapons production. In the studies cited, inspectors detected from 30 to 90% of existing defects, depending upon many variables. Lower defect detection percentages are associated with low defect rates (1% or less), passive inspection, and inspection for several types of defects in each unit. The highest detection percentages are associated with higher

Table 15-1. Estimated Probabilities of a Checker's Failure to Detect Errors

Checking Operation	HEP
Usual type of monitoring in NPP with some kind of checklist or written procedure (includes tasks such as over-the-shoulder checking and checking written lists or procedures)	.10 (.05 to .5)
Same as above but without written materials	.20 (.10 to .9)
Special short-term, one-of-a-kind checking (e.g., supervisor checks performance of a novice)	.05 (.01 to .10)
Hands-on type of checking that involves special measurements or other activities	.01 (.005 to 0.5)
Repeated checking of one person's work by different individuals during or after completion of a standard or routine task	Assume HD among checkers; no recovery credit for more than 2 checkers
Two-man team performs task, with one person the doer and the other the reader/checker	Assume HD between doer and reader/checker
Checking of valve status	Use HEPs from Ch. 13
Walk-around checking operation	Use HEPs from Ch. 8

defect probabilities, more active participation in the production process, and inspection for only one or a very few well-defined defects.

Inspection and checking in NPPs are generally not as passive as in typical industrial assembly tasks, and the kinds of signals the checker is looking for are usually well-defined. In WASH-1400 we assigned a 10% error probability to the monitoring or checking of another's activities and have found no data to warrant modification of that basic HEP as a general approximation. However, there are cases in which the type of inspection is "one of a kind"; e.g., when an operator asks someone else to check something. This constitutes a break in the general work procedures, and the checker can be expected to approach the task with a higher level of alertness and attention. For such nonroutine monitoring tasks, we divided the basic HEP of .10 by 2, for an HEP of .05. For routine monitoring performed without written materials, we have doubled the basic HEP to .20 and have assigned pessimistic error bounds because errors of oversight are especially likely without recourse to written materials. The behaviors of an operator and a monitor are not independent; if the monitor believes that the operator's work is reliable, he tends to assume that the operator's performance will be correct. This assumption and the resultant perceptual set or expectancy reduce the checker's effectiveness; he is likely to miss an operator's error because he does not expect it. Even when the error is clearly visible and involves no interpretation, the checker will often fail to "see" it.

There are cases in which the opposite influence between an operator and a checker occurs; for example, when the operator is a novice the checker may take extra care in his checking routine. This relationship cannot be predicted in the abstract, but if such a case is known to exist when performing a human reliability analysis, the analyst may wish to regard the job of the checker as more active, and the .05 HEP can be applied to the checker.

In some checking tasks, the job of the checker is not passive and may involve hands-on activities such as using an instrument for some measurement. In such cases we regard the performance level of the checker as approaching that of the original operator. The tabled value of .01 applies to this kind of an active monitoring task that includes much more involvement than is usual with a checker.

The use of several checkers in succession to check someone's performance has limitations because the second, third, and later checkers do not expect to find anything wrong and may see this type of assignment as "make-work." There are no studies directly related to the loss of efficiency in repeated checking activities in an NPP. In one experiment with electronics assembly plant personnel (Harris and Chaney, 1969, pp 79-80), 45% of the defects were detected in the first inspection, with about 15% more defects found in each of the next two inspections. After the third inspection, further inspections were much less effective, and after the sixth inspection essentially no further defects were found. It would be inappropriate to generalize these results without modification to the task of checking in NPP operations. The above experiment was artificial in that the subjects knew it was an experiment and saw it as a challenge (the subjects were experienced inspectors). In the NPP work situation, we believe that the motivation of the second and subsequent checkers would not facilitate effective checking (except for abnormal circumstances which all recognize as important). Although we have designated HD as the level of dependence to assume between successive checkers, we recommend that no recovery credit be allowed for more than two checkers for a routine task. For analysis, then, use an HEP of .5 for the second checker. (Walk-around inspections are a separate case, as discussed in Chapter 8.)

Sometimes, as when performing calibration procedures, two men act as a team, with one man responding to the other's verbal instructions, which are read from a set of written instructions. The [illegible] also acts as checker. Because the checker's task is passive, we estimate a high level of dependence in estimating the reader/checker's HEP. For analysis, assume an HEP of .5 for the reader/checker. The checking of valve status is different, and is discussed in Chapter 13.

There is a misconception about performance that is frequently expressed; the belief that an operator's HEP will be substantially increased if he knows that his work will be checked, and that greater reliability might be achieved without a checker. From the HEPs assigned to checkers it is clear that such a state of affairs would rarely occur, if ever. The increase in the HEP of the operator would have to be greater than the overall increase in reliability resulting from the recovery factor of the checker. We do not know of any such cases in the type of work addressed in this handbook.

The estimates in Table 15-1 apply to checking tasks that are actually performed. The following section presents some estimates of failure to perform checking tasks.

Administrative Control

Administrative control refers to the checking of human performance mandated in a plant and the extent to which this and other plant policies are carried out and monitored, including the use of inventory systems. Good administrative control means that certain types of errors can be held to a minimum. For example, as noted in chapters 13 and 14, lower HEPs are associated with proper use of written materials and with proper use of tagging and locking systems. Good administrative control increases the likelihood of proper use of these important job aids.

The quality of administrative control can be inferred from the extent to which written procedures and checklists are properly used, the type and use of tagging and locking systems for critical valves, the type of inventory system and its use, and the general attitudes and practices of operating personnel. These PSFs are plant-specific, ranging from poor to good at different plants.

Following are some examples of poor administrative control that make errors more likely than if good control is used. In one plant the use of step-by-step written procedures was mandated by plant management with severe penalties, such as loss of pay, for offenders. Yet we frequently observed that the highly skilled technicians performing the work made only casual reference to the procedures. Much of the time, in fact, the page to which the written procedures were opened did not list the operations being performed. This problem was reduced only when the procedures were rewritten so as to remove excess wordiness, yielding a higher signal-to-noise ratio (as described in Chapter 14).

At several plants, we have observed people working without mandated checklists or using them incorrectly; e.g., performing several steps and then checking them off all at once on the checklist. This same casual attitude has been noted in the use of tags and keys. At one plant, keys for critical valves were distributed to several operators. It cannot be said that the keys were administratively controlled.

A final example shows the importance of evaluating the extent to which mandated human redundancy will be used properly. This example comes from an assembly plant where great emphasis was placed on the quality of the product and minor emphasis on the quantity of product. At certain stages of assembly in one system, the assembler was instructed by the written procedures to call for an "over-the-shoulder" inspection of a torquing operation. At this point, the inspector was supposed to consult his own procedures to determine the value of torque to be applied. Then he was to look over the shoulder of the assembler while the latter applied torque by means of a torque wrench with a built-in meter. When the assembler was finished, the inspector was to indicate if he agreed with the assembler regarding the torque that had been applied.

If carried out correctly, the above procedure is estimated to result in a 90% recovery factor; i.e., one time in ten the inspector would fail to note an incorrect torque setting. However, the manner in which the procedure was actually carried out removed all human redundancy and amounted to just one person's performing without any checking. First, the assembler called to his friend, the inspector, "Hey, Joe, I need an over-the-shoulder." Joe replied, "What's the torque supposed to be?" The assembler informed Joe from his (the assembler's) procedures. Then he proceeded to the torquing operation with Joe looking over his shoulder. The assembler did not even look at the meter; he looked back at Joe. Eventually, Joe said, "OK, that's it," and the assembler relaxed the torque wrench. If a human reliability analyst improperly allowed the 90% credit normally assigned to this recovery factor, the joint probability of failure would be underestimated by a factor of 10.

Clearly, in performing a human reliability analysis some estimate should be made of the percentage of people who carry out the various procedures properly and the percentage of those who do not. If one is performing a reliability analysis of plants in general, without knowledge of the types of administrative control at the plant, it is still necessary to form some estimate of the extent of good or poor administrative control. Based on our experience at several plants, we list the estimated HEPs for these activities in Table 15-2. We believe these estimates are conservative. Since there is so much uncertainty in an across-plants analysis, the conservative estimates are intended to ensure that a human reliability analysis is not overly optimistic.

Table 15-2. Estimated HEPs Related to Failure of Administrative Control

Operations	HEP
Carry out a plant policy when there is no check on a person	.01 (.005 to .05)
Initiate a checking function	.001 (.0005 to .005)
Use control room written procedures under operating conditions that are	
Normal	.01 (.005 to .05)
Abnormal	no basis for estimate
Use a valve restoration list	.01 (.005 to .05)
Use written calibration procedures	.05 (.01 to .1)
Use written maintenance procedures when available	.3 (.05 to .9)
Use a checklist properly (i.e., perform one step and check it off before proceeding to next step)	.5 (.1 to .9)*

---

*Some reviewers regard this as an overly optimistic estimate.

Tagging Systems

In most plants a tagging system is used to indicate nonnormal status of equipment; e.g., a normally closed valve has been opened to permit testing or maintenance. The use of tags and their management is one of the most important methods for ensuring awareness of nonnormal equipment status and the later restoration of that equipment to normal status. As used here, the term tagging system includes all administrative controls that ensure (1) awareness of any valves or other items of equipment that are in a nonnormal state, and (2) prompt restoration of this equipment to the normal state after the completion of test or maintenance operations. Thus, a tagging system includes the use of tags; chains, locks, and keys; and logs, suspense forms, and any other techniques that provide an inventory of the above items.

Obviously, the quality of tagging systems can vary widely. As a guide for human reliability analysis, we identify three levels of tagging systems, listed in Table 15-3. Since a system of locks and keys is considered as part of the tagging system, Table 15-3 includes three levels of lock and key control. In human reliability analyses, we do not differentiate between tags per se and locks and keys. If both are used, one identifies the highest level of control and bases and the analysis on that level. To be conservative, no extra credit should be allowed where both exist, even when both are the highest level.

Unless otherwise stated, the HEPs in the handbook are premised on a Level 2 tagging system. If tags are not used on items changed from their normal operating state, this practice suggests extremely poor administrative control. As a gross approximation of the HEPs in such a "system" for tasks normally associated with tagging, multiply each relevant HEP by 10, with a maximum HEP of .5.

Table 15-3. The Three Levels of Tagging Systems

Level	Tags
1	A specific number of tags is issued for each job. Each tag is numbered or otherwise uniquely identified. A record is kept of each tag. In addition, a record of each tag issued is entered in a suspense sheet that indicates the expected time of return of the tag; this suspense sheet is checked each shift by the shift supervisor.
2	Tags are not accounted for individually - the operator may take an unspecified number and use them as required. In such a case, the number of tags in his possession does not provide any clues as to the number of items remaining to be tagged.
3	Tags are used, but recordkeeping is inadequate to provide the shift supervisor with positive knowledge of every valve, circuit breaker, or other item that is in a nonnormal state.

Level	Locks and Keys*
1	The number of keys is carefully restricted and under direct control of the shift supervisor; a signout board is used for the keys; keys in use are tagged out; and each incoming shift supervisor takes an inventory of the keys.
2	The shift supervisor retains control of the keys and records the issuance of keys, but does not use visual aids such as signout boards or tags.
3	Keys are generally available to users without logging requirements.

---

*Locks and Keys are considered part of an NPP's Tagging System.

The rating of tagging systems must be based on a thorough evaluation of plant practices, as there is a great deal of inconsistancy in the application of tagging practice even within individual plants. For example, at one plant we observed a Level 1 tagging system. All tags were numerically assigned, logged, and accounted for by a dedicated operator designated the Tagging Controller. Before maintenance, the Tagging Controller would tag the appropriate items of equipment and retain a tab stub in a special file. The other stub would be issued to the maintenance organization. After maintenance the maintainer would return the stub(s) for the completed job to the Tagging Controller who would then go to the appropriate items of equipment and remove the tags for which he had been given stubs. If there were no other tags on those items of equipment, he would restore them. On returning to his office, the Tagging Controller made the necessary entries in his log. He could also tell from the log if any item of equipment was still tagged for pending tests or maintenance.

This is clearly an excellent tagging system. However, on occasion this excellent system is circumvented. Sometimes the Tagging Controller is too busy to restore some valves, and he removes the tags from a set of valves but leaves the valves in the nonrestored positions. He then reports to the shift supervisor or to the senior control room operator, requesting him to assign another operator to complete restoration by using a valve lineup sheet. The fact that tags are removed before valves are restored creates opportunities for serious oversights due to any number of reasons. Thus, the reliability of the normally excellent tagging control system breaks down.

Another serious problem occurs at the same plant during shutdown, when a great number of valves are placed in nonnormal positions. No tagging is used at shutdown because of the very large number that would be required.

Realignment of valves is accomplished with valve alignment lists. A single person is responsible for the restoration of any given valve -- no human redundancy is utilized. Such reliance on an individual in the case of valve restoration is typical in NPPs, and is not unique to this particular plant.

In evaluating a tagging system, one has to estimate which level is being used, what percentage of time that level is used, and the likelihood of errors of omission or commission in preparing tags. The probability of each type of error is estimated as .003 (.001 to .01) (p 14-7). Finally, in estimating the probabilities of human errors, the use or nonuse of human redundancy must be noted. Even if one assumes a probability of .10 for failure of a checker to note an unacceptable condition, failure to check at all means that the possibility of an order-of-magnitude reduction in the overall failure probability has been lost.

CHAPTER 16. DISTRIBUTION OF HUMAN PERFORMANCE AND UNCERTAINTY BOUNDS

This chapter treats two related topics:

(1) We present some background data on the distribution of human performance in a variety of settings and offer some hypotheses about the distributions of HEPs in NPP tasks.

(2) We present guidelines for establishing uncertainty bounds around point estimates of HEPs in NPP tasks. The estimated uncertainty bounds are wider than the presumed variability among NPP personnel so as to include other sources of uncertainty when assigning the nominal HEP for a task.

Data and Hypotheses About Human Variability

All human performance displays variability within the same individual and among individuals. Generally intra-individual variability is small compared with the variability among different persons in an NPP. Moreover, it is difficult (if not impossible) to predict a specific individual's day-to-day variability. In this handbook the concern is with inter-individual variability -- the variability among properly trained and experienced operating personnel. The estimated uncertainty bounds in the handbook include both sources of variability.

Despite variability among people, managerial influences in most technical jobs tend to restrict variability under normal operating conditions. If a person consistently performs far below the average for his group, he usually will be retrained, reassigned, or terminated. If he performs in a consistently exemplary manner, he usually will be promoted or transferred to a more challenging and responsible position (Peter and Hull, 1969).

The Normal Distribution

Most human traits and abilities do not conform to the Gaussian (normal) distribution. As pointed out by Wechsler (1952), the only human distributions that approximate the normal curve are those which pertain to the linear measurements of people, such as stature, lengths of extremities, the various diameters of the skull, and certain ratios like the cephalic index. Even among these there is often a considerable departure from the symmetry of the normal distribution.

Wechsler, a psychologist well known for his development of scales to measure adult intelligence, measured the abilities of many people in such disparate tasks as running the 100-yard dash and assembling radio tubes in a factory. He noted that the usual measure of human variability, the standard deviation, was highly sensitive to the shape of the distribution. He therefore suggested a different measure of variability, which he called the total range ratio. He defined this as the ratio of the highest score to the lowest score of a group of people homogeneous with respect to what is being measured,* but excluding the extreme scores; i.e., the lowest and highest tenths of 1% of the population (Wechsler, 1952, p 46). He discovered that when the extreme scores were discarded, the ratio of the scores of the best performers to the scores of the worst performers was generally on the order of 3:1 to 5:1. He further discovered that if only production tasks were considered, the typical range ratio was rarely over 3:1 and more typically approached 2:1 (Table 16-1). He noted that this ratio was probably influenced by unwritten

---

*By homogeneous, Wechsler means that all the individuals considered were subject to the factors which produced the characteristics that influence what is being measured. Therefore, one would not compare persons in running speed when only some of them were especially trained in this ability or when some of them had some physical impairment.

Table 16-1. Range Ratio in Production Operations*

Operation	Unit or Criterion	Subjects (Adults)	Mean	S.D.	Extremes	Ratio**
Filament mounting (elec. bulbs)	Avg no. per hour	100	104.9	13.5	132-75	1.73:1
Assembling radio tubes	Number of tubes done per hour	65	99.5	16.9	132-61	2.00:1
Manipulating automatic lathe	Avg output per day	120	71.8	6.85	104-51	2.04:1
Card punching	Avg no. cards punched per hour	113 on day shift	232.2	29.8	340-160	2.10:1
Checking, posting, and listing	Percent of avg	34 females	102.1	20.7	150-65	2.30:1
Checking grocery orders	Time in seconds	46	365.0	--	575-225	2.53:1
Machine sewing	Units per hour (5 day avg)	101 females	72.4	14.3	112-43	2.55:1
Hosiery looping	Dozen per hour	99 experienced loopers	4.0***	--	7.2-2.8	2.57:1
Typing	Number of words per minute	616	53.4	9.57	85-30	2.83:1

---

*From Wechsler, 1952, p 68

**The ratio is computed by discarding the extreme values in the previous column.

***Median

agreements among workers. This peer restriction of performance has long been recognized in production work, even with piecework systems, in which workers are paid in direct relation to their output.

Wechsler's data pertain to quantity of output, not generally a concern in NPP operations. Wechsler states, "Of the two basic characteristics of production, namely, amount and accuracy, only that of amount can be treated unequivocally in studying range ... If errors (or accuracy) alone are taken as a base, they are likely to prove misleading." (p 65)

To study the validity of this limitation, L. W. Rook, then of Sandia National Laboratories, analyzed 6 months of production defect data from a large industrial plant that manufactures electronic and electromechanical components. (This unpublished 1962 study is described briefly in Swain, 1980b, pp 64 and 69.) Rook had the following hypothesis:

> "In an industry where production is more important than errors, the production process will tend to weed out those workers who produce far less than the average worker and will tend to promote those who produce far more; this process would tend to generate range ratios of quantity of production of less than 3:1. But in an industry such as the nuclear weapons industry where production quantity is not nearly so important as the number of defects, the production process will tend to weed out those workers who generate more defects than the average worker and will tend to promote those who generate far fewer; this process would tend to generate a range ratio of defect frequencies of less than 3:1."

When Rook tabulated individual differences for each of several types of defects, he found range ratios approximating 2:1 for numbers of defects, so he felt justified in saying, "It has been observed many times that, in typical situations, the best workers are almost never more than three times better than the worst workers. In fact, a ratio of two to one is more typical" (Rook, 1965). This ratio, known as Wechsler's Range Ratio, can be stated as follows:

> Given a population that is homogeneous with respect to some ability that can be measured with a ratio scale, if the extreme scores are discarded (i.e., the lowest and highest tenths of 1%), the highest score in the distribution will rarely be more than five times the smallest score, and usually not more than three times the smallest score.

Based on the above, if we have a good estimate of the mean of some ability, and if we assume a roughly normal curve, we can construct an estimated distribution by assigning the appropriate range ratio to the ±3 standard deviation (SD) points. For example, if we assume a range ratio of 3:1, the ±3 SD points in the distribution correspond to values 50% greater and less than the mean. The lower bound would be the point estimate of the HEP divided by 2 and the upper bound the HEP times 1.5. We have used this method to derive estimated extremes of performance in human reliability analysis in weapon systems.

The Lognormal Distribution

As noted earlier, often the assumption of a normal curve is not valid. When data are collected in experiments of human performance, different shapes and types of distributions occur. Normal, lognormal, Weibull, Gamma, exponential, or other distributions may closely fit the empirical distributions (see Green and Bourne, 1972, Chapter 7, for a discussion of various distributions).

We believe that the distribution of the logarithms of HEPs for NPP tasks is often normal (or approximately so), even though the distribution of the

HEPs per se is not; i.e., the HEPs are lognormally distributed. Our rationale is that performance of skilled persons tends to bunch up towards the low HEPs. This is quite unlike the case in the typical laboratory experiment in which the task is made deliberately difficult so that a high HEP can be obtained for statistical convenience. In such cases, one is much more likely to obtain a curve approximating the normal distribution of individual differences. However, in the typical industrial setting, especially with highly skilled performers, most scores fall near the low end of the error distribution.

There are data that support the use of a lognormal distribution for the performance of skilled people. In one study, an analysis of human performance data revealed lognormal type distributions for simple tasks and slightly skewed distributions approaching the normal for more complicated tasks (Swain, 1967b). A lognormal distribution was reported in a British study of the time taken to respond to a simulated alarm signal superimposed on normal tasks in an NPP (Green, 1969). In an unpublished followup study in Danish research reactors, similar results were found by Jens Rasmussen and his coworkers at the Risø National Laboratory.

The parameters of the applicable lognormal distribution are, of course, speculative. We hypothesize that for most NPP tasks a lognormal probability density function (pdf) with an SD of 0.42 would provide a suitable fit (Figure 16-1). This SD of 0.42 was obtained by assuming a 4:1 range ratio between the 95th and 5th percentile on the dimension of error probabilities. Until more data can be collected on the performance of NPP personnel, such hypothesized curves are merely suggestive.

Relative Unimportance of Assumed Distributions

Although we would like to have data clearly showing the distributions of human performance for various NPP tasks, there is ample evidence that the

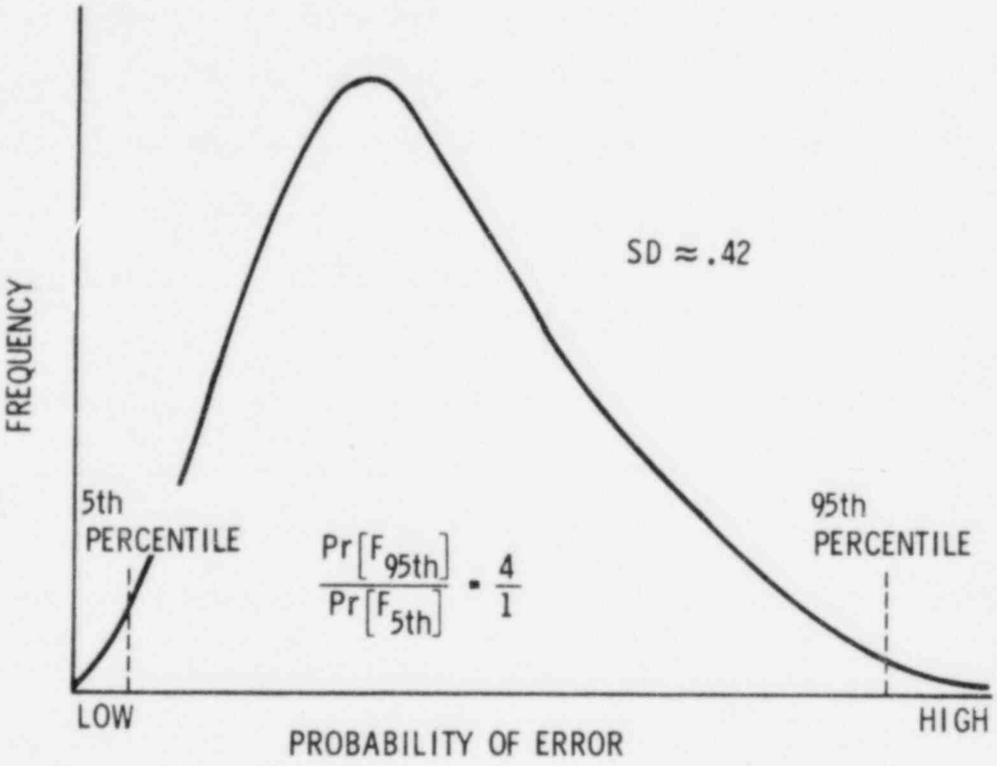


Figure 16-1. Hypothesized Lognormal Probability Density Function for Nuclear Power Plant Personnel

outcomes of human reliability analyses are relatively insensitive to assumptions about such distributions. One example was described in the appendix to Chapter 8.

In an experimental study, Mills and Hatfield (1974) collected distribution data on task completion times for several realistic tasks. One significant finding was that errors resulting from assuming an incorrect form of the pdf of task times were small and of little practical significance. The distributions, all unimodal, were best described by various Weibull functions, but the authors state, "From a practical point of view, ... the error produced by assuming normality may not be large enough to warrant using a better-fitting pdf."

During the preparation of WASH-1400, a sensitivity analysis was performed by Dr. W. E. Vesely using a Monte Carlo procedure to see if the assumption of different kinds of pdfs for the human failure estimates would materially affect the availability of various subsystems in safety-related systems. It was found that the predicted unavailability did not differ materially no matter what distribution was assumed.

We can conclude that, for human reliability analysis of NPP operations, the assumption of normal, lognormal, or other similar distributions usually will make no material difference in the results of the analysis. In some cases, this insensitivity may result from a well-designed system which has so many recovery factors that the effect of any one human error on the system is not substantial. However, if some very different distributions such as the exponential or extreme value were used, it is possible that different results could be obtained. For computational convenience one might wish to assume the same distribution for human failure as the one used for equipment failure. A sensitivity analysis will reveal whether any significant differences will be obtained with different assumptions.

Uncertainty Bounds

Generally the use of single-point estimates of task reliability are adequate for reliability analysis work. In such work we simply ignore uncertainty in the estimate. However, distributions must be obtained or assumed for certain applications, as in WASH-1400 or the bounding analyses described in Chapter 21 (p 21-20).

The preceding sections indicate the range of variability in performance that may be expected under routine performance of well-defined tasks such as factory production work. In estimating the range of variability to be expected in NPPs, we have to make allowances for the greater variety of tasks, for the less routine nature of much of the work, and for a large number of unknowns, such as the nature of the plant, the relevant PSFs, administrative practices, and other such variables discussed in other chapters. Consequently, the range ratios used in reliability analyses of NPP tasks will be considerably wider than the nominal 4:1 ratio described in Figure 16-1, since there is uncertainty in the nominal HEPs assigned to the tasks as well as in the variability of the tasks. In consideration of the many sources of uncertainty, we have extended the uncertainty bounds around the estimated HEPs to limits that we feel confident include the true HEPs. For example, for an estimated HEP of .01 (.005 to .05) the numbers in parentheses represent a lower and an upper bound of uncertainty. These bounds reflect our judgment regarding the likelihood of various values of HEPs. We believe that the bounds include most of the range of HEPs resulting from individual differences and other unspecified sources of uncertainty. It is difficult to associate an exact probability statement with the bounds. However, it is our intention that the lower bound correspond to the 5th percentile and the upper bound the 95th percentile of this distribution regarding HEPs. If the user of this handbook has better estimates of the distribution of performance for some given application, he should use them. In the meantime, Table 16-2 can be consulted for some general guidelines in assigning uncertainty bounds. The uncertainty bounds in the table are generic; in several cases, the uncertainty bounds we have assigned to tasks in the handbook are less than the generic bounds in the table. In some situations the user may wish to assign even larger uncertainty bounds than those in the table.

The relation of the estimated uncertainty bounds to the nominal HEPs varies among tasks. For some tasks, the divisor of the nominal HEP used to obtain the lower bound (LB) is smaller than the multiplier used to obtain the upper bound (UB). This would be the case when, due to human variability and our knowledge of the task, the upper bound HEP is much "farther away" from the nominal HEP than is the lower bound. For example, the estimated nominal HEP is .01 for improperly mating a connector (last item in Table 12-1, p 12-7). The estimated lower bound is .005 (a divisor of 2) and the upper bound is .05 (a multiplier of 5). For other tasks it might be judged that the lower bound HEP should be "farther away" from the nominal HEP than is the upper bound. If this were true in the above case, we might have used a divisor of 5 and a multiplier of 2, yielding lower and upper bounds of, respectively, .002 and .02. Until distribution data on task HEPs become available, the user will have to use either our or his judgment in assigning uncertainty bounds.

To account for the modifying effects of certain PSFs, the nominal HEP is multiplied or divided by some factor. This factor is also applied to the uncertainty bounds, with upper limits as noted.

We have assigned different divisors and multipliers for the uncertainty bounds for nominal HEPs $< .5$ and $\geq .5$. For example, in Table 11-4 (p 11-21) we use the following rule (see p 16-12):

Table 16-2. General Guidelines for Estimating Uncertainty Bounds for Estimated HEPs*

Task and/or HEP Guidelines	Uncertainty Bounds**	
	Lower	Upper
HEP-Oriented		
Estimated HEP < .001	HEP ÷ 10,	HEP x 10
Estimated HEP .001 to .01	HEP ÷ 3	HEP x 3
Estimated HEP > .01	HEP ÷ 5	HEP x 2 to x 5
Task-Oriented		
Task consists of step-by-step procedure conducted under routine circumstances and essentially static (e.g., calibration task), HEP $\geq$ .001	HEP ÷ 3 to ÷ 5	HEP x 3 to x 5
Same as above, but estimated HEP < .001	HEP ÷ 10	HEP x 10
Task consists of step-by-step procedure, but carried out in nonroutine circumstances specifically involving a potential turbine/reactor trip	HEP ÷ 5	HEP x 10
Task consists of relatively dynamic interplay between operator and system indications, but task is done under routine conditions; e.g., increasing or reducing power	HEP ÷ 10	HEP x 10
Task performed under severe stress conditions; e.g., large LOCA, conditions in which the status of ESFs is not perfectly clear, or conditions in which the initial operator responses have proved to be inadequate and now severe time pressure is felt	.03	.75

---

*The estimates in this table apply to experienced personnel. The performance of novices is discussed in Chapter 18.

**The lowest lower bound is 5×10^{-5} and the highest upper bound is .999.

For HEPs $< .5$,

$$\mathrm{LB} = \mathrm{HEP} \times 0.2$$

$$\mathrm{UB} = \mathrm{HEP} \times 2$$

For HEPs $\geq .5$,

$$\mathrm{LB} = 1 - 2(1 - \mathrm{HEP})$$

$$\mathrm{UB} = 1 - 0.2(1 - \mathrm{HEP})$$

Reference to Table 16-2 shows that the large uncertainty bounds for HEPs smaller than .001 reflect the greater uncertainties associated with infrequently occurring events. The cutoff for the lower bound at 5×10^{-5} is based on data collected in weapons production where the lowest HEP for very small units of behavior (e.g., omitting a component from a circuit board in a well-designed assembly process) was 3×10^{-5} (Rook, 1962; and Rigby and Swain, 1968). The units of behavior in the handbook are larger, and 5×10^{-5} is used as the lowest lower bound.

HEPs in the range of .001 to .01 generally apply to routine tasks involving rule-based behavior. HEPs greater than .01 are uncommon and are associated with tasks performed under conditions conducive to error, such as performance under high stress levels or checking the status of items that provide no indication of their correct status. Uncommon cases should be evaluated individually to determine appropriate error bounds, as we have done in developing the error bounds for many tasks cited throughout the handbook. Often, it is appropriate to use nonsymmetric bounding with more allowance for error probabilities larger than the estimated HEP.

CHAPTER 17. STRESS

Chapter 3 (pp 3-53 to 3-65) presents a brief discussion of the performance shaping factors psychological and physiological stress. This chapter extends that material and presents HEP estimates and uncertainty bounds for different levels of stress.

Objective data on the effects of stress are spotty, and there is no comprehensive treatment of the effects of stress on performance available at this time. In particular, there are very little data on the performance of technical personnel under stress in an applied setting. In this chapter we attempt to apply what little is known about stress to the performance of NPP personnel. Some of these extrapolations are made with confidence, others are frankly speculative, especially for situations in which rule-based behavior does not apply.

In Chapter 3 we said stress results from a mismatch between the external and internal PSFs acting on an individual. An everyday definition of stress would be "any situation that causes tension within the individual." Although we usually think of tension as an undesirable state, we will demonstrate that a certain amount of tension can be beneficial.

The Four Levels of Stress

The classical stress curve in Figure 17-1 indicates that stress and performance have a curvilinear relationship and that stress is a continuum ranging from very low to extremely high. For human reliability analysis, it is adequate to represent the entire continuum of stress by the four levels in the figure:

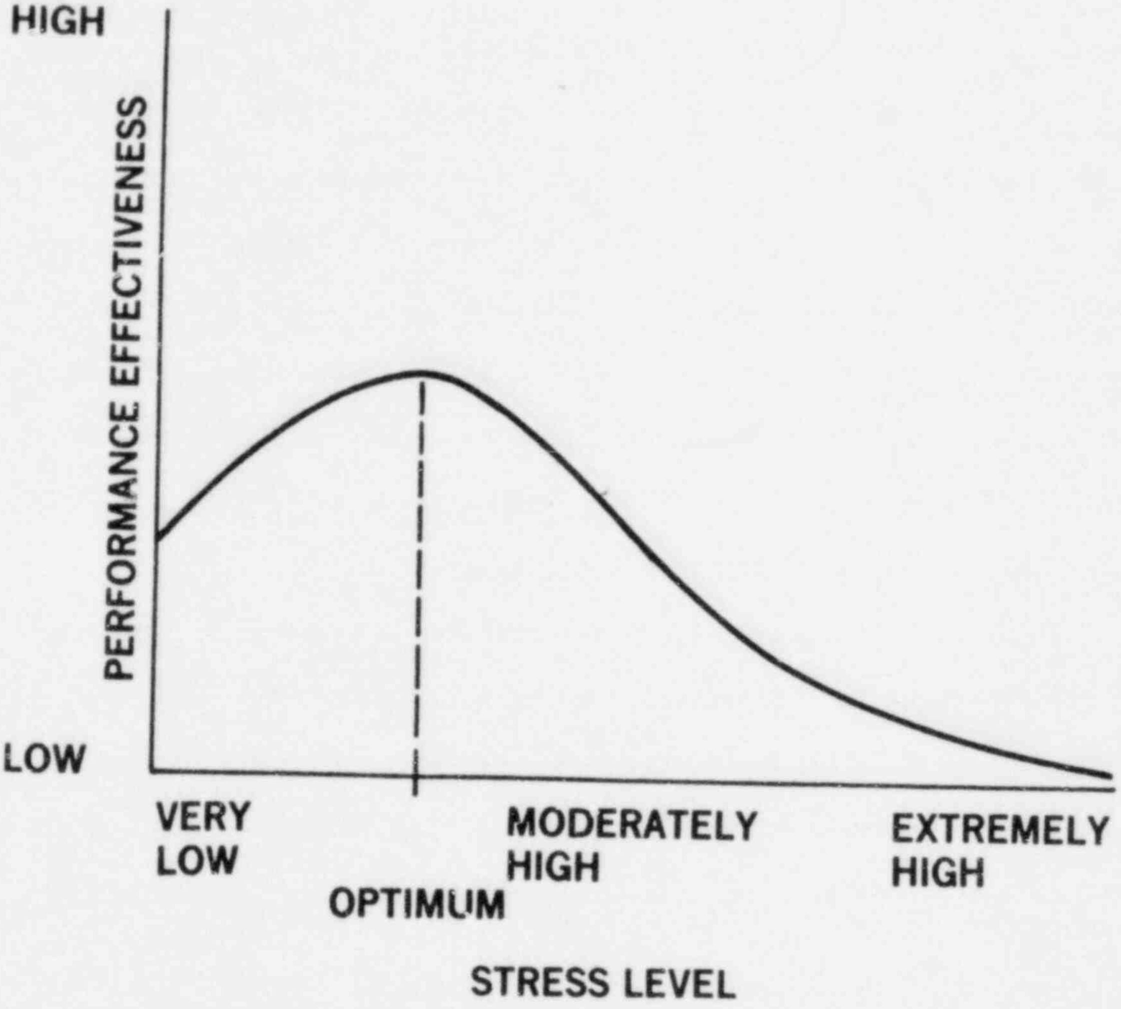


Figure 17-1. Hypothetical Relationship between Performance and Stress (based on Figure III 6-1 from WASH-1400)

(1) Very low (insufficient arousal to keep alert)

(2) Optimum (the nominal or facilitative level of stress)

(3) Moderately high (slightly to moderately disruptive for human behavior)

(4) Extremely high (very disruptive for most people)

The effects of the first three levels of stress can be approximated by applying modifying factors to the HEPs and uncertainty bounds in the handbook. The fourth level of stress is qualitatively different from the other three levels; i.e., the effects of this level of stress will outweigh most of the other PSFs. For this reason, a different set of HEPs and uncertainty bounds is assigned to the extremely high stress level situation.

Since stress is a response to some stressor(s), the curve in Figure 17-1 is only a gross approximation of reactions to situations that <u>most</u> people consider stressful. There is great variability in perceived stress - a situation that seems threatening to a novice may be perceived as routine by a more experienced person. For example, a novice driver feels considerable stress when entering a busy freeway, whereas an experienced driver perceives it as a commonplace occurrence. The stress curve is intended to represent the relationship of performance to the perceived level of stress, and the perception of stress will vary with an individual's knowledge, experience, preparation, personality, and many other factors. The curve represents average performance for a large number of people. The uncertainty bounds assigned to the HEPs at different levels of stress are intended to include the middle 90% of individual differences in response to stress.

Unless otherwise stated, the HEPs in the handbook presume an optimum level of stress. In some cases, the rationale for an estimate will be based on other than the optimum stress level; e.g., the high HEP estimate in Chapter 8 for the relatively passive task and low arousal of the walk-around inspection. In performing a human reliability analysis, it is necessary to decide whether the stress level for a task is other than optimum, and, if so, to modify the HEPs accordingly.

The rest of this chapter presents discussions of each level of stress and guidance in determining the stress levels associated with various tasks and conditions in NPPs.

Very Low Stress Level

The decrease in performance effectiveness of vigilance tasks may be attributed to a decline in the person's level of arousal, or alertness, caused by a lack of stimuli.* The lack of sufficient stimuli is the condition that we describe as the very low stress level. It is a familiar phenomenon that, as people have less to do, they tend to become less alert. As this period of very low stress persists, the level of alertness decreases even further.

For the very low stress level, most people will manifest a minimal level of alertness. Some control room tasks, such as the periodic scanning of unannunciated displays described in Chapter 11, are charac-

---

*There is controversy in the psychological literature over the whole concept of vigilance, especially on the applicability of laboratory studies to industrial settings (Buckner and McGrath, 1963, Jerison and Pickett, 1963, O'Hanlon and McGrath, 1968, Smith and Lucaccini, 1969, Craig and Colquhoun, 1975, and Mackie, 1977). However, Fox (1975) cites studies which show evidence of the vigilance effect in industry. In this handbook, we use vigilance merely as a descriptive term.

terized by such a low level of arousal. For periodic scanning tasks, the HEP estimates have taken this very low level of stress into account. Arousal can occur very quickly, as when an auditory signal sounds. For this reason, a separate set of data was developed for the responses to auditory annunciators in Chapter 10.

Lower error probabilities are estimated for tasks with a specific requirement to look at a particular display at a particular time. This reflects the higher level of arousal generated by a requirement to do a specific task rather than merely to look around the control room periodically to see if everything is as it should be. In NPPs, tasks involving detection of infrequent signals are the most likely to suffer performance degradation due to low arousal levels.

As a working rule, we suggest that the estimated HEPs and uncertainty bounds pertaining to tasks performed under the optimum level of stress be multiplied by a factor of 2 if the tasks must be performed under conditions of very low arousal (or stress). To illustrate this modification, assume the .003 general HEP for an error of omission or commission for a task performed under optimum stress. Also assume that HEP ÷ 3 and HEP x 3 estimate the lower and upper uncertainty bounds, yielding .001 and .01 (rounded). The ratio between the 95th percentile error probability and 5th percentile error probability is .01/.001 = 10:1. Now assume that this task is to be performed under the very low stress level. Using the factor of 2, the new HEP is .006 with uncertainty bounds of .002 and .02. Thus, the same 10:1 ratio between the extremes of the middle 90% of the HEPs is maintained.

A final point: overqualified persons are more likely to experience a very low level of stress than less qualified persons for whom the tasks would prove interesting or challenging. Traffic safety studies reveal that very bright people make poor cab drivers -- the job is too boring and they tend to daydream, incurring a disproportionate number of accidents. Similarly, telephone answering services have found that employee turnover was highest among those who had the highest levels of education - again a matter of inadequate challenge (or arousal) in the task. Industrial literature has many articles on the deleterious effects of assigning a job to a person whose qualifications are far in excess of the job demands. (For a short human factors review of this point, see Swain, 1973.) If the analyst judges that the person assigned to a given job in an NPP is so overqualified that the job would be regarded by him as dull and uninteresting, the very low level of stress, and its modifying factor of 2, should be assigned. Of course, the occasional assignment of some uninspiring tasks to a highly qualified person does not fall into this category. The key issue is whether the person perceives his overall job as dull and boring.

At the opposite end of the continuum is the person with limited capacity, who finds challenge in a job that most would consider dull. The utilization of slightly retarded people in simple industrial jobs, such as elevator operations, has been very successful. The applicability of this practice to selected NPP jobs could be investigated.

Optimum Stress Level

Unless otherwise stated, the HEPs in this handbook are based on a level of stress that we judge to be optimal for most people. This is the optimum stress level in Figure 17-1, and is characterized by an active interaction between the person and his environment -- talking with others, reading displays, adjusting controls, making decisions, etc, at a pace that the person can manage comfortably.

Examples of tasks for which we assumed an optimum stress level are: maintenance and calibration, the initial audit of the control room, the reading of an annunciated legend light, the decision to initiate some action in response to an annunciator, and the scheduled reading and recording of some quantitative value from a display.

Moderately High Level of Stress

Most people operating under a moderately high level of stress experience some degradation in their performances. Such disruption could be caused by a requirement to perform at a faster pace than the person is capable of or a requirement to make prompt decisions in a situation in which a wrong decision could result in costly consequences. Because of the obvious subjectivity in assessing a moderately high level of stress in a reliability analysis, it is difficult to state which NPP tasks fall into this category. We adopt a conservative rule of assigning this level of stress to transients that involve shutdown of the reactor and turbine (but excluding large LOCAs), to certain tasks during startup and shutdown which must be performed within time constraints, and to work performed in a radiation environment where special protective clothing must be worn. We can offer no detailed guidelines in making this judgment; it tends to be situation- and plant-specific. In general, if it is judged that most personnel will feel a good deal of time-pressure without an accompanying high level of emotional stress, the moderately high level of stress should be presumed.

A LOCA is a special case. Presumably a small or slowly developing LOCA should not be accompanied by more than a moderately high level of stress for most people. Of course, in some incidents involving small LOCAs, the initial stress level may not be very high, but subsequent events may raise the stress level. For example, some operating personnel in the TMI accident, which involved a small LOCA, were considered subject to high levels of stress at various times by the interviewers on the Kemeny Commission (Kemeny, 1979) and the Rogovin Special Inquiry Group (Rogovin and Frampton, 1980).

We can find no objective data from which to derive the factors to apply to HEPs and uncertainty bounds for the condition of a moderately high level of stress. On the basis of judgment, we multiply the HEP and uncertainty bounds for step-by-step, rule-based tasks performed under optimal stress levels by 2, and for tasks requiring dynamic interplay between the operator and system indications, we use a multiplier of 5. Thus, if one takes the .003 general HEP for a task done under optimum stress levels, with uncertainty bounds of .001 and .01, under moderately high stress the HEP will be .006 (.002 to .02) for step-by-step tasks and .015 (.005 to .05) for dynamic tasks.

For the moderately high level of stress, the level of arousal is so high that the net end effect is disruptive. For transients and other unusual events requiring quick response, a wider distribution of operator performance may be expected than for normal plant operations. Some operators will respond to the unusual with a calm, cool approach to solving the problem. Other operators may "freeze," mentally withdraw from the situation, or even panic. The diversity of reaction is a function of many PSFs. Perhaps the three most important are (1) the emotional stability of the operator, (2) his skill level associated with the unusual condition in question, and (3) the extent to which displayed information directly supports the actions the operator should take to cope with the situation.

Regarding the first PSF, we cannot predict the emotional stability of the operator. Although there are tests of emotional stability (Matarazzo, 1972), none have been validated in the NPP environment.

The second critical PSF is the operator's skill in responding to the unusual situation. A problem here is that, apart from an operator's formal training, which includes practice of some simulated transients and emergencies on dynamic simulators, he rarely receives further practice after he is assigned to an NPP. Thus, his skill level for coping with unusual events is expected to follow the skill decay curve in Figure 3-15 (p 3-69). Our general estimate of HEPs under moderately high stress levels takes into account this lack of specific practice. If personnel at a plant indeed have such frequent practice that the tasks in question could be regarded as "second nature," the HEPs assigned to the moderately high level of stress will not apply, as the stress level will be closer to optimum. In judging whether plant personnel have the necessary skills, one must determine that they really do receive the frequent practice required. Some NPP personnel we have interviewed believe that an operator has a high state of skill in a task as long as he performed that task at the last training session in a dynamic simulator, even though that session might have occurred many months ago and there has been no practice since then. We strongly disagree with such an assessment.

Chapter 21 (p 21-14) presents an example of a human reliability analysis in which the use of in-plant practice justifies an HEP estimate of essentially zero for carrying out a well-rehearsed sequence of actions, given that the correct decision has been made to initiate the sequence.

The third critical PSF is related to the inadequate human engineering of equipment. According to the Rogovin Report, as long ago as 1975 Dr. Stephen Hanauer of the NRC said, "Present designs in NPPs do not make adequate provisions for the limitations of people" (Rogovin and Frampton, 1980). This report also states, "During the period in which most large nuclear plants have been designed, the nuclear industry has paid remarkably little attention to one of the best tools available for integrating the nuclear operator into the system: the relatively new discipline of 'human factors.'" -- It continues, "The NRC gives short shrift in the design safety review process to determining how well operators will be able to diagnose abnormal events, based on what they see on their instruments, and respond to them." In the TMI accident, these limitations were manifested by the operators' incorrect and tardy diagnoses, which worsened an already serious situation. It seems that much of the stress experienced by the operator in an emergency may be due to his inability to diagnose the cause of the emergency. In other words, the displays do not present all the <u>essential</u> data in an immediately usable form, nor do they help the operator filter the essential from the nonessential data. The inability to "size up the situation" promptly is, in itself, a stressful experience, as has been observed in simulator exercises. We intended that our conservative estimates of HEPs take these instrumentation limitations into account.

The above three variables are related. Military experience shows that a person with a tendency to panic or freeze in an abnormal situation is much less likely to do so if he is thoroughly skilled in diagnosing and responding to the situation and if the information he is provided is directly related to the problem at hand. The high HEPs and large uncertainty bounds for performance under stress may seem pessimistic, but they are justified in view of the human factors inadequacies in existing plants.

Extremely High Level of Stress

The Data Problem

Most of the experimentation on high levels of stress deals with artificial tasks in situations in which the experimental subjects clearly realize that nothing catastrophic will result from any ineptitude on their parts. See Harris et al (1956), Klier and Linskey (1960), and Robbins et al (1961) for literature reviews. A subsequent search of the literature indicates that this situation has not changed (Bell, 1980).

Another body of literature deals with the performance of military personnel under combat stress (Grinker and Spiegel, 1963; and Marshall, 1978). The latter reference reports results of interviews with World War II combat soldiers in which it was found that an average of not more than 15% of the men interviewed had actually fired at enemy positions or personnel with rifles, carbines, grenades, bazookas, Browning automatic rifles, or machine guns during an entire engagement (p 54). This suggests that most of the men were ineffective, assuming that the remaining 85% had opportunity to fire their weapons. In the best companies no more than 25% of the personnel used their available firepower even though most of the actions occurred under conditions in which it would have been possible for at least 80% of the men to fire, and in which nearly all personnel were at some time operating within firing distance of the enemy. This suggests about 30% net effectiveness (i.e., .25 ÷ .80).

Comparisons of the stress resulting from a large LOCA or its equivalent in an NPP with the stress of combat are obviously open to error. First, in combat, the participant's life is literally and obviously at stake, and he perceives this as fact. Such is not the case in most NPP emergencies. Second, death in combat is seen and is not rare; close calls are numerous. In NPPs, real life-threatening emergencies are rare. Third, much of the evidence on combat stress includes the effects of combat fatigue (i.e., the cumulative effects of unrelieved stress over a long period of time) and therefore does not apply to an industrial situation. Fourth, combat training emphasizes coping with emergencies; this is a major purpose of training. In NPP training, responding to emergencies is only a small part of the training, since nearly all of the tasks to be performed are routine and are performed in a relatively stress-free situation.

Despite the risks in generalizing the results of military studies on stress to the behavior of NPP personnel, we will consider two such studies which are classics in the applied area of stress. In one series of studies, performance of soldiers was measured under conditions in which they did not realize that the experimental stressors were artificial. They really thought that either their own lives were in danger or that they had caused others' lives to be endangered (Berkun et al, 1962; and Berkun, 1964). In the other study, critical incidents were collected on U.S. Air Force aircrews who survived in-flight emergencies (Ronan, 1953).

Because our estimates of HEPs under very high levels of stress are based on these two studies, they are described briefly. The major problem in estimating the performance of NPP personnel under very high stress is that very few of them have been subjected to this level of stress. Moreover, the kind of accident that is generally considered to represent the highest level of stress in an NPP, a large LOCA, has never occurred and is unlikely to occur. The best we can do is generalize from other types of emergency situations that may be only marginally related to the performance of NPP personnel under very high stress.

The Ronan Study

In the Ronan study, aircrews surviving in-flight emergencies on the B-50 (a propeller-driven aircraft with four reciprocating engines) were interviewed and critical incidents were noted. The critical incident technique (CIT) was used by U.S. Army Air Corps investigators in World War II and has subsequently seen extensive use in postwar military investigations and in safety analyses (O'Shell and Bird, 1969). As defined by Flanagan (1954), "The critical incident technique (CIT) consists of a set of procedures for collecting direct observations of human behavior in such a way as to facilitate their potential usefulness in solving practical problems and developing broad psychological principles. Fivars and Gosnell (1966) define an incident as "... any observable bit of human behavior sufficiently complete in itself to permit inferences to be made about the person performing the act." They further state that for an incident to be critical, "... an incident must make a significant difference in the outcome of the behaviors; it must contribute either positively [a positive incident] or negatively [a negative incident] to the accomplishment of the aim of the activity." Traditionally, such a negative contribution has been called a "red" incident and a positive contribution a "blue" incident.

In the Ronan study, aircrews from several Air Force bases were interviewed, using a carefully structured interview with assurance that all reports should be strictly confidential; i.e., no word as to who said what would get back to the interviewee's superior officer. Dr. Ronan stated that there was no apparent reluctance on the part of the interviewees to describe in detail the positive and negative critical human behaviors they observed (or engaged in). The data we use are from 153 aircraft commanders (ACs). These ACs were highly trained, with an average of 2971 flying hours. Twenty-nine categories of emergencies were described, ranging from very serious emergencies such as engine loss on takeoff to less serious problems such as an engine with rough operation. We have excluded one category, crew coordination problems, since we wished to base the derived HEPs only on AC performance. The ACs submitted a total of 2450 critical incidents over the remaining 28 categories. Of these, 360, or 15% of the total, were red incidents. Thus, of the critical actions taken during in-flight emergencies, 15% were ineffective in that the situation was not improved or made worse as a result of them. This percentage probably underestimates all the errors made by ACs, since nonsurvivors could not be interviewed.

Not all the emergencies resulted in reports of ineffective behavior. Considering the entire aircrew, there were 1229 reported emergencies and 457 red incidents, an average of .37 red incidents per emergency. Thus, in some emergency situations there were no red incidents (the exact data are not presented in the Ronan report).

In WASH-1400, we equated the percentage of red critical incidents with the error probability for tasks performed by ACs under the stress of in-flight emergency situations. (In WASH-1400, we cited the median of the 28 ratios of red to total incidents, 16%, rather than the simple mean of 15%. In view of the nature of the data, these numbers can be considered equivalent.) Obviously, the 15% figure is not the same kind of error probability defined in Chapter 2 (the number of incorrect responses divided by the number of opportunities for response). In the Ronan study, the denominator is not the number of opportunities, but the number of red and blue critical incidents reported. Therefore, there is some unknown error in considering the 15% figure as an error probability.

The Berkun Studies

In the Berkun studies, raw recruits and experienced soldiers were placed in several elaborately simulated emergency situations. Data from the few who perceived the deception were excluded from the results reported. A Subjective Self-Report showed that those subjected to the various experimental (stressful) conditions reported they felt "timid," "unsteady," "nervous," "worried," or "unsafe," whereas the control subjects (those subjected to nonstressful conditions) reported that the conditions "didn't bother me," or that they themselves were "indifferent" to the situation. Interview results indicate that the experimental group of subjects really believed that their safety was endangered or that their actions had imperiled the safety of others.

In the Berkun ditching study, subjects were passengers on an apparently stricken aircraft that was forced to crash-land in the ocean. Through a plausible fabrication of events, the passengers were led to fill in two forms, one midway through the flight and the other close to the presumed ditching point. That only 5 of the 20 experimental subjects saw through the deception illustrates the realism of the study. The performance scores on filling in the forms for the remaining 15 were significantly lower than for a flying control group, showing a 10% decrement for the first form administered and a 36% decrement for the second form. Presumably, the stress level for the second form in the experimental group was higher since the aircraft was very close to the presumed ditching location in the ocean. In terms of error probabilities, the experimental subjects had an HEP of .29 compared with .21 for the control subjects, on the first form, an increase of 38%. On the second form, the HEPs for the experimental and control subject were .59 and .37, respectively, showing a 59% increase for the experimental group.

In three other test situations, the subject was led to believe that he was in immediate danger of losing his life or of being seriously injured. The performance task required him to "repair" a defective radio to summon help. The situations differed only in the events contrived to cause the simulated emergency: accidental nuclear radiation in the area, a sudden forest fire in the area, or misdirected incoming artillery shells. In these studies, the only statistically significant differences in performance scores between the control subjects and the experimental subjects were in the artillery study. Here there was a 33% decrement for the latter subjects, largely because 8 of the 24 experimental subjects panicked and fled the scene as the simulated artillery shells came closer. In this study, the HEPs were 39.5% for the control subjects and 59.5% for the experimental subjects.

Derivation of the Estimated HEP Under Extremely High Stress on NPP Tasks

In WASH-1400, we used data from the Ronan study and the Berkun artillery study to establish boundaries for the error probabilities of NPP personnel in a high stress situation such as a large LOCA. It is not likely that NPP personnel would react as calmly or perform as reliably as Air Force aircrews, considering the extensive practice of such crews in simulated aircraft emergencies. Generally, there is little if any onsite practice in simulated emergencies for operators of nuclear power plants. Hence, the error probability of plant operators was estimated to be higher than the .15 error probabilities observed by Ronan.

On the other hand, NPP operators have been extensively trained in NPP operation. A newly assigned operator has learned to recognize the potential for accidents and to be prepared to cope with them, although not at a level comparable with that of the Air Force personnel. Hence, we judged that the operators should be better able to cope with emergencies than would raw Army recruits, and their error probability should be lower than the .33 probability of (completely) inadequate behavior of the artillery subjects. Therefore, we assumed an estimate of .25 for the average error probability for NPP personnel in a high stress situation such as a large LOCA. This estimate is based on the further assumption that the perceived stress in a large LOCA situation compares to the perceived stress in the Ronan and Berkun studies, whereas in fact it might be lower.

It is regrettable that there are no better data available to develop HEPs for the high stress condition in NPPs. Until better data are collected, this kind of rationale is all we have.

Uncertainty Bounds for Very High Stress

Using the .25 HEP as the estimate of the central tendency of trained NPP personnel under very high stress, it is necessary to assign uncertainty bounds for some reliability calculations. In the absence of actual data, the following interim uncertainty bounds for errors of omission or commission are offered. For the 5th percentile, we assign an estimated HEP of .03. This estimate is based on our judgment that under very high stress the HEP would be about a factor of 10 higher than it would be under optimum stress, which is represented by the general HEP of .003. The .03 permits 5% of our estimated HEPs to indicate more reliable performance, even exemplary performance.

For performance under extremely high stress, we judge that an HEP of about .75 represents the 95th percentile HEP. Using an HEP of .75 yields a range ratio of 25:1 between the upper and lower bounds. Thus, the estimated range ratio under very high stress is 2.5 times the largest range ratio (10:1) assumed for tasks performed under optimum stress levels.

The Stress PSF and the Large LOCA

In WASH-1400, we used a large LOCA as the example of a situation resulting in a very high stress level for the operators, and we estimated the HEPs for an operator from the first moments of a large LOCA until the operating crew could establish control of the situation. Figure 17-2 shows our estimates as a function of time after the onset of the accident. The curve is speculative since a large LOCA has never occurred. Our rationale for the curve has not changed since WASH-1400, and was then explained as follows:

> "Following a LOCA, human reliability would be low, not only because of the stress involved, but also because of a probable incredulity response. Among the operating personnel the probability of occurrence of a large LOCA is believed to be so low that, for some moments, a potential response would likely be to disbelieve panel indications. Under such conditions it is estimated that no action at all might be taken for at least one minute and that if any action is taken it would likely be inappropriate.
>
> "With regard to the performance curve, in the study the general error [probability] was assessed to be .9 five minutes after a large LOCA, to .1 after thirty minutes, and to .01 after several hours. It is estimated that by seven days after a large LOCA there would be a complete recovery to a normal, steady-state condition and that normal error [probabilities] for individual behavior would apply."
>
> (WASH-1400, p III-61)

The solid line in Figure 17-2 indicates that the above estimates from WASH-1400 apply if the automatic recovery systems function normally to mitigate the effects of the accident. Otherwise, as indicated by

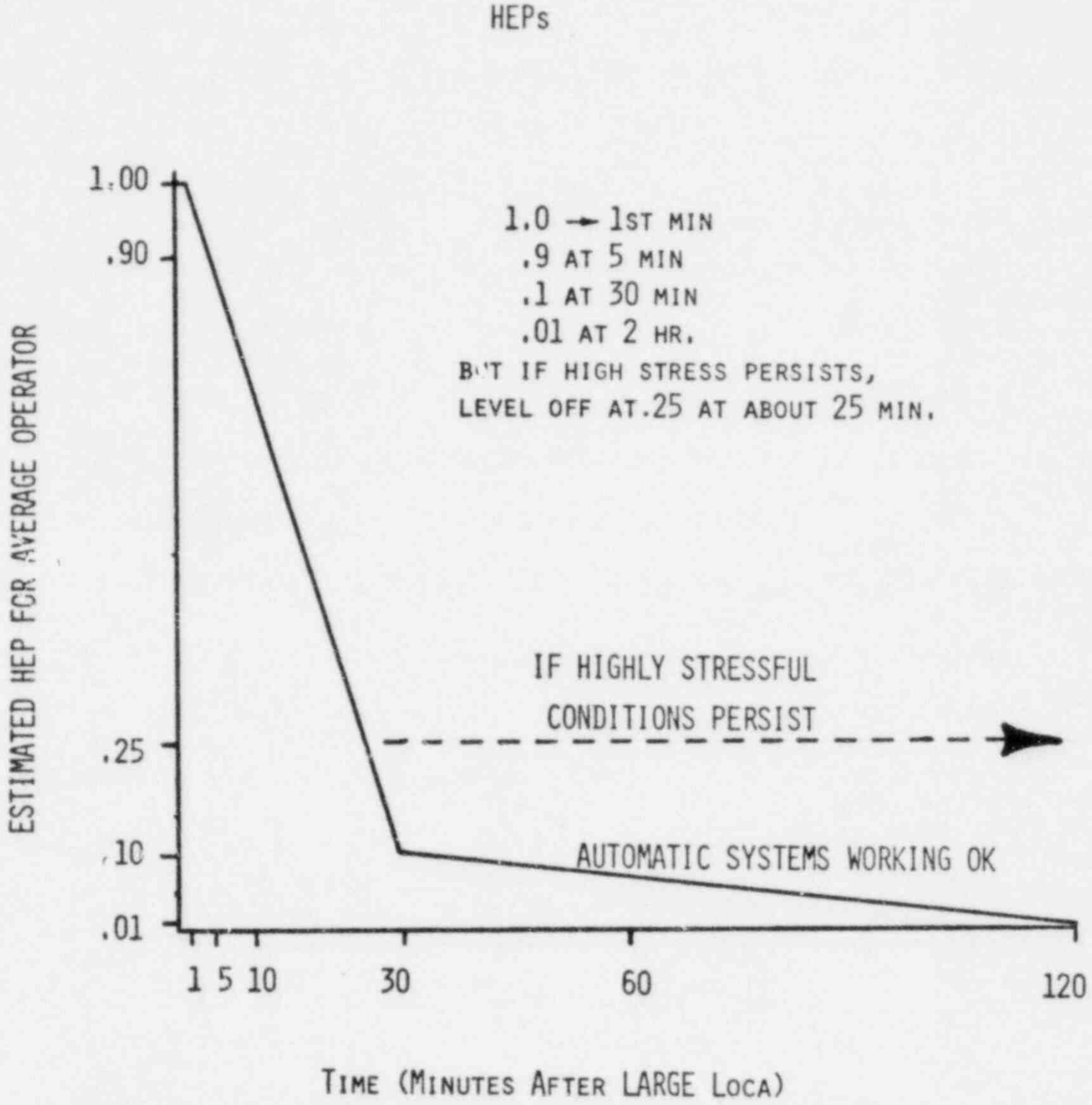


Figure 17-2. Estimated Human Performance after a Large LOCA

the dotted line, the error probability will not decrease below .25 but will remain at that value as long as the highly stressful conditions persist. The large uncertainty bounds around the .25 estimate (.03 to .75) allow for some individuals to perform well and for others "to be a part of the problem."

It has come to our attention that some people have misapplied the large LOCA curve to LOCAs of all types. Such an application is inappropriate for a very small LOCA, as in most cases people will have enough time to diagnose the problem and take appropriate action. A moderately high level of stress would usually be more appropriate.

The Doubling Rule

There is an important exception to the shape of the performance curve in Figure 17-2 as well as to the estimated HEPs for the condition of moderately high stress. This exception applies if operators are required to take corrective action in response to a LOCA (or other stressful situation) and the time available to take this corrective action is severely restricted. In our early work on human reliability analysis (Swain, 1963b), we developed a theory of human behavior under time stress. This theory holds that, given that an initial error has been made and perceived as such or that an initial action has failed to have its intended corrective effect, the error probability for each succeeding corrective action doubles. Thus, if one starts out with an error probability of .2 and fails on the first attempt at corrective action, it takes only three more unsuccessful attempts to reach the limiting case of an error probability of 1.0. This limiting condition corresponds to the complete disorganization of an individual. Extensive clinical experience indicates that large numbers of individuals will fail to perform assigned tasks under severe stress and may become completely disorganized (Appley and Trumbull, 1967; Grinker and Spiegel, 1963; and Marshall, 1[illegible]8). Experimental studies of Naval aircraft pilots landing on carrier decks (Siegel and Wolf, 1969), indicate a stress curve with repeated attempts after failure that closely matches the doubling rule. Although the doubling rule was developed for very high stress conditions, we now believe that it is valid for moderately high stress conditions also.

An obvious analog to the doubling rule is the halving rule, which states that, under optimum stress conditions, a person takes extra care once he has made a mistake, and his HEP on his next attempt is half his nominal HEP for the task. If we could identify all the important PSFs and possible modes of human behavior for the performance of a task even under optimum conditions, we would feel justified in using the halving rule. With our present state of knowledge, we offer the halving rule as interesting speculation only. We do not use it in our reliability analyses as it could lead to undue optimism about human reliability.

Effects of Several Operators in an Abnormal Situation

In WASH-1400 (p III-68) we used an HEP of .1 for the tasks involved in changing over from the injection to the recirculation mode after a large LOCA. The .1 was premised on an estimate that 30 minutes after the LOCA these procedures would be performed, by which time the average HEP would be down to .1 (Figure 17-2). We further assumed that when these tasks were attempted there would be at least three qualified operators in the control room and that their joint HEP would be approximately 10^{-3}. At first glance this 10^{-3} seems to presume ZD among the three operators; i.e., $(10^{-1})^3$. Our rationale was that, although the assumption of ZD might be inappropriate, the 10^{-1} estimate was gross enough that cubing this HEP for three people would not materially affect the results of the analysis.

We now believe this estimate may have been too optimistic; but, as shown in Chapter 21 (p 21-1), the more conservative approach described below does not change the estimated impact of human errors on the tasks evaluated. Sheridan (1980), in writing about the TMI accident, states,

> "Nuclear plant operators work in teams, based on the premise that two or more heads are better than one. But there is a great deal of interaction among team members, some of it subtle and unspoken. Such interpersonal communication is little understood but assuredly does affect the reliability of human performance. For example, operators unintentionally could reinforce one another's misimpressions, making the team less reliable than a single operator who would be more likely to think a matter through carefully. This means that human error rates for individuals may differ from those for teams."

In the Rogovin Report, one can see that it took some time for available operating personnel to become organized to the extent that people were stationed at strategic places rather than running around (Rogovin and Frampton, 1980).

Until data can be gathered on the performance of more than one operator in an abnormal situation, we offer the following performance rules. For the performance of any given task in an abnormal event such as a transient.

(1) During the first 5 minutes following the initiation of the event, assume only one operator is present at the control room panels for a reactor. This assumption is based on interviews with operating personnel and on statements from shift supervisors that the shift supervisor could be as many as 5 minutes away from the control room. This is obviously a worst-case situation because usually the shift supervisor is only a minute or less away. Five minutes is selected as a maximum on the assumption that the plant security system would not hold up the shift supervisor if he happened to be outside the guarded area. If such an assumption is not reasonable for a given plant, the 5-minute estimate should be increased accordingly. After 5 minutes into the abnormal event, assume that the shift supervisor is present.

(2) Depending on the task to be evaluated, assume a moderate-to-high level of dependence between the shift supervisor and the assigned operator. Thus, the joint probability of failure of the two to take some appropriate action will be the product of the BHEP (modified by the factor for the appropriate level of stress) and a conditional probability of .15 or higher (Table 7-2, p 7-34).

(3) After 20 minutes into the event, assume the presence of another qualified operator for assistance in coping with the event. Depending on the task, assume a high-to-complete level of

dependence between this person and the others. The assumption of a 20-minute delay is made for the worst-case situation in which there is a minimum crew (as on the night shift) and there is another reactor which must be kept running. The assumption of high-to-complete dependence is based on two additional assumptions: (1) the third person may be an auxiliary operator or reactor operator with limited experience, and (2) there may be a stressing effect because of the "Tower of Babel" influence mentioned by Sheridan (1980) (see p 17-23).

(4) After 2 hours into the event, assume a conservative maximum of three qualified personnel to be present, as we assumed in WASH-1400. (At some indeterminate time there will be more qualified people present; however, we know of no way to assess their influence.) Beyond 2 hours into the event, the presence of additional qualified people may or may not help cope with the event. The Rogovin Report describes instances in the TMI accident when erroneous diagnoses were still being made beyond 2 hours into that event even though several additional personnel were present.

The BHEPs assigned must take into account the skill levels of the people involved in coping with the assumed event. For example, most operators receive considerable training and simulator practice in coping with a large LOCA before being qualified as reactor operators, but receive very little training or practice thereafter. If the personnel in a given plant maintain high levels of skill in coping with a large LOCA, the HEPs from the curve in Figure 17-2 could be reduced considerably, in some cases by a factor of 10 or more. However, for analyses that are not plant-specific, use the HEPs from Figure 17-2 without modification. Chapter 21 compares the analysis in WASH-1400 for the probability of correctly shifting from the injection to the recirculation mode with an analysis using the above performance rules. The calculations in Chapter 21 show that if the above rules are used, the end result is essentially the same as that obtained using the original assumptions.

CHAPTER 18. SKILL LEVEL

The skill levels of qualified NPP personnel can be represented by a continuum ranging from acceptable to superior. (As discussed in Chapter 16, anyone whose skills were less than acceptable would not be retained.) To facilitate reliability analysis, we specify only two levels of skill: experienced and novice. We arbitrarily define the novice as a person with less than 6 months on the job for which he has been licensed (in the case of reactor operators, operating control rooms) or otherwise qualified (in the case of auxiliary operators, maintainers, and technicians).

Assumptions

The models and estimated HEPs in this handbook are based on the assumption that all NPP personnel are experienced people who have been trained and qualified for their jobs rather than novices. The uncertainty bounds for the HEPs allow for individual differences in skill among the experienced personnel.

This chapter presents modifying factors for novices. Table 18-1 presents modifying factors to apply to the tabled values of HEPs (and their uncertainty bounds) for novices and experienced personnel under different levels of stress and for different types of tasks. Although some data (Berkun, 1964; and Berkun et al, 1962) indicate that experienced people perform more reliably than novices under highly stressful conditions, we are taking the conservative position that the difference between their performances is negligible. Therefore, the effects of very low and very high stress are assumed to be the same for novices and experienced personnel. Under optimum stress, the novice is assumed to

Table 18-1. Comparison of Estimated HEPs and Uncertainty Bounds of Experienced Personnel* and Novices** Under Different Levels of Stress

	Type of Task			
	Step-by-Step Procedures		Dynamic Interaction	
Stress Level	Experienced	Novice	Experienced	Novice
Very low	Nominal x 2	Nominal x 2	N/A	N/A
Optimum	Nominal	Nominal	Nominal	Nominal x 2
Mod. High	Nominal x 2	Nominal x 4	Nominal x 5	Nominal x 10
Extremely High	.25 (.03 to .75)	.25 (.03 to .75)	.25 (.03 to .75)	.25 (.03 to .75)

---

*See Chapter 20 for the tables of nominal HEPs and uncertainty bounds for experienced personnel which are used in this table as the basis for comparison.

**A novice is defined as a person with less than 6 months on the job in which he has been licensed or otherwise qualified.

be twice as error-likely when performing tasks involving dynamic interaction (which includes decision-making). The greatest differences between novices and experienced personnel are expected to develop under conditions of moderate stress, as when responding to transients. The modifying factors in the table are based on our judgment, and may be changed as data become available.

Control Room Operators

In the case of operators in the control room, Seminara et al (1976) state, "...some shifts are manned entirely with novices, with only several months' experience." This statement refers to the manning of the control room, primarily in the evening or night shift, and does not refer to the shift supervisor, who must be a highly experienced senior reactor operator. We judge that, for nonroutine tasks, the HEPs for the novices will be about twice that of experienced operators. In view of the requirements for becoming a licensed reactor operator (3 years of NPP experience including 1 year at the plant where he is licensed, of which 6 months shall include experience as an auxiliary operator), it might seem that there is no difference in performance between an experienced operator and a novice operator. Our distinction is based on the following rationale: the 3 years of training prior to licensing is certainly adequate for acquisition of the knowledge required to operate the plant; however, about 6 months of unsupervised experience is required for a person to develop the confidence to exercise his decision-making authority fully when he is responsible for the resolution of some unusual problem that might arise.

Over a 6-month interval, the new operator will experience one or more transients and will develop a "feel" for the plant which cannot be acquired in trainee status. In most industrial settings, 6 months is accepted as the time required for a person to achieve full performance capability after a promotion or reassignment. Thus, although 6 months was stated as an arbitrary interval for distinguishing between novice and experienced operators, it is based on observation in comparable industrial settings and on the opinions of senior reactor operators

The factor of 2 difference between the HEPs of novice and experienced control room operators would have most of its effect in the first 5 minutes after the occurrence of an abnormal event because, for a worst-case analysis, one would estimate that the shift supervisor would not be available (p 17-24). After the shift supervisor arrives, the estimated HEPs for tasks should be based on his performance rather than on that of the novice.

Other Personnel

For auxiliary operators, maintainers, and technicians the same factor of 2 in estimated HEPs is used between the performances of novices and experienced personnel. This factor of 2 may be an underestimate for maintenance novices. Based on observations and interviews by Seminara, Parsons et al (1979) and by Brune and Weinstein (1980a), it seems that the training of maintenance personnel is not as complete as that of auxiliary operators and technicians. Most of the maintainer's training of NPP specifics is on-the-job, with considerable dependence on his previous background. Therefore, it is possible that the difference between a novice and an experienced maintainer may be greater than that reflected by the factor of 2. However, with no objective data, the factor of 2 will be used as an interim figure.

PART IV. AN INTERIM HUMAN PERFORMANCE DATA BANK

PART IV. AN INTERIM HUMAN PERFORMANCE DATA BANK

This part of the handbook consists of two chapters: Chapter 19, "Sources of Human Error Probability Data," and Chapter 20, "Derived Human Error Probabilities and Related Performance Shaping Factors."

Chapter 19 describes some of the studies that were used to derive the estimated HEPs in the handbook and comments on the shortage of objective data on human performance in NPPs. Chapter 20 summarizes the HEPs in Part III of the handbook. After one has used the handbook for several human reliability analyses, most problems will require reference only to the tables in Chapter 20.

CHAPTER 19. SOURCES OF HUMAN ERROR PROBABILITY DATA

This chapter describes some of the sources of HEP estimates used throughout the handbook and summarized in Chapter 20. Background data which were presented elsewhere in the handbook are not repeated.

Categories of the Sources Used

The error probabilities in this handbook are extrapolated from a variety of sources, and the nature of these sources determines the confidence we have in these derived HEPs. The following source categories were drawn on in gathering the error probabilities.

1. Nuclear power plants
2. Dynamic simulators of NPPs
3. Process industries
4. Job situations in other industries and in military situations that are psychologically similar to NPP tasks (For example, errors in radar maintenance were used in estimating some maintenance errors in NPPs, and errors in measurement tasks in the nuclear weapons field were used in estimating calibration errors in NPPs.)
5. Experiments and field studies using real-world tasks of interest; e.g., experimental comparisons of measurement techniques in an industrial setting, performance records of industrial workers, etc
6. Experiments using artificial tasks; e.g., typical university psychology studies which have limited application to real-world tasks

The above listing orders the sources by their relevance to NPP tasks. Unfortunately, the availability of HEP data is just about the inverse of this order. Available human performance data relevant to estimates of HEPs for NPP tasks are discussed below for each of the categories.

Nuclear Power Plants

Hardly any HEPs for NPP tasks have been recorded. In WASH-1400, we used some data collected by the United Kingdom Atomic Energy Authority (Ablitt, 1969). To date there has been no systematic program to collect HEP data in operating NPPs in the U.S. The only formal record of errors in NPPs consists of the Licensee Event Reports (LERs), which do not yield probability data in the sense of errors per opportunity for error. In some cases it may seem possible to derive a probability from an LER. For example, if an error is made on a task and it is known that the task is performed, say, once a day every day, we have the denominator for that task, and it seems possible to calculate an HEP for that task. However, the HEP obtained will be the unrecovered error probability, which will have to be adjusted for the unknown number of times that an error is made which does not result in a reportable event; e.g., when recovery factors compensate for the error. This number should constitute the actual numerator of the HEP.

In one study (Joos et al, 1979), "gross human error probabilities" were derived by counting the number of each type of error reported in the LERs submitted by operating U.S. commercial NPPs over a 25-month period. These numbers were then divided by the total number of months that the plants had been operating to arrive at "the number of errors per plant month." Although such numbers do indicate the relative frequency of different types of reported events, they do not yield the type of error probabilities required for a human reliability analysis. In most real-world situations, recovery factors will prevent a single error from resulting in an event. We have no way of estimating appropriate factors by which to adjust LER data to obtain estimates of the basic error probabilities.

Dynamic Simulators of NPPs

Dynamic simulators have been used extensively--almost exclusively--for training and recertifying NPP operators. Although there have been no extensive programs to gather HEP data, the simulators are readily adaptable for this function, and the Electric Power Research Institute (EPRI) has plans to gather error probability data in conjunction with training programs. Although a simulator is not "the real thing," it is expected that HEPs obtained from advanced trainees and from experienced operators undergoing recertification will be very valuable.

Swain (1967a) has described a technique for combining real-world data with simulator data in such a way that real-world data serve as "anchor points," permitting valid extrapolations of the simulator data to the real-world situation.

Process Industries

The operation of process industries such as chemical plants, refineries, etc, is very similar to that of NPPs. Some data have been collected in various process industries, many of which are directly applicable to NPP operations (e.g., Kletz, 1970, Kletz and Whitaker, 1973; and Edwards and Lees, 1974). Extensive use has been made of this data, in some cases with modifications to allow for differences in task details.

Industrial and Military Data

Many job situations in the nuclear weapons industry, in the military procedures for the handling of weapons, and in other military tasks are similar to tasks performed in an NPP. In some of these areas reliable error data were available which were directly applicable to NPP tasks.

Field Studies

Field studies and experiments in industrial settings yielded some reliable data directly applicable to NPPs. A problem with many experimental procedures is that the very fact that an experiment is being conducted tends to influence the performance of the workers. Some allowance has to be made for this effect. Data gathered in field studies are less susceptible to this effect, but these data are harder to obtain and less complete.

Experiments Using Artificial Tasks

Data obtained in university laboratories usually deal with artificial tasks but are very precise, because all pertinent variables can be tightly controlled and there is ample time to gather a large amount of data. The studies carried out in university laboratories usually answer specific academic questions, and the findings may be applicable to real-world situations after appropriate modifications are made.

A feature of academic studies is that the limits of acceptable performance are often very narrow. Also, the situation is often arranged to be artificially difficult so that the distribution of responses will be "forced" into some desired form (usually the normal distribution). The reported performance of subjects under these conditions could lead to a pessimistic evaluation of human capabilities in real-world situations, so the unmodified data are not used. Allowances are made for the broader tolerances in an industrial situation. As an example, Payne and Altman (1962) multiplied HEPs from laboratory experiments by a factor of .008 to derive HEPs for certain maintenance tasks.

Many of the academic experiments are comparative studies to determine how basic performance is affected by different conditions. In such studies we can often use the comparative performance data directly, since we are interested in the ratio of performance under the differing conditions rather than in the absolute levels of performance.

Expert Opinion

All the data sources listed above yield numbers based on some kind of documented records of performance. In addition to error estimates based on such "hard" data, estimates are prepared on the basis of expert judgment. For example, in those instances in which error terms from a particular task are modified to apply to some task in an NPP, the modification is necessarily based on a judgment of the similarities and differences between the tasks and the extent to which these would affect the error probabilities. The uncertainty of the judgments is reflected in the extent and direction of the bounds assigned to the error estimates.

Error probability modifications based on judgment may be made informally or formally. Informal evaluations of error estimates are usually obtained from just a few (two or three) experts who are thoroughly familiar with the tasks and the relevant PSFs. Their opinions are elicited informally and pooled to arrive at the extrapolated error probability. Formal evaluations require a larger number of judges (up to 10) and involve standardized evaluation forms. The judges may not have the intimate task knowledge required of those who make informal judgments. The data are analyzed by statistical methods known as psychological scaling techniques (Kendall, 1948; Coombs, 1952; Guilford, 1954; Remmers, 1954; Morrissey, 1955; Edwards, 1957; Torgerson, 1958; Thurstone, 1959; Gulliksen and Messick, 1960; Blanchard et al, 1966; and Swain, 1967a, 1977b). Although the formal judges may not have the thorough task knowledge that the informal judges have, their pooled evaluations can provide useful estimates of error probability modifications.

In this handbook, when using expert judgment, we have relied on informal judges because of their greater familiarity with the tasks in question and their ready availability. In many instances, because of the paucity of relevant "hard" data, judgments were the only source of error probability estimates. In all such cases the judgments were based on data from tasks that most nearly resemble the task in question, and the error tolerances were adjusted in accordance with the judged similarities or differences between the tasks.

For many years, the HEP of .01 was routinely assigned to human actions by reliability analysts in weapon systems. For lack of more specific data, the figure of .01 was used for almost all individual actions identifiable as "units of activity." It was recognized that .01 was often a pessimistic estimate of HEPs, but this conservative estimate was used in view of the nature of the systems. In applying human reliability analyses to other fields, more realistic of HEPs were desired, and the figures in the handbook were developed for application to the specific actions with which they are associated.

To the reader not familiar with human reliability analysis, the HEPs in the handbook may seem to be too conservative. It must be stressed that these values represent the probability that the error will be made at all, <u>not</u> the probability that the error will remain uncorrected. For example, let us say that there is an HEP of .01 that a person will attempt to insert a key into a lock upside down. Thus, if a person starts his car three times a day, about once a month he will insert his key in the ignition switch incorrectly. This does not mean that the error will be uncorrected. In this particular situation the error is always self-corrected, yet the basic HEP is .01 (a fairly high probability). Similarly, the HEPs presented throughout this handbook are the basic HEPs; i.e., the probability that the error will occur at all. The effects of relevant PSFs and recovery factors must be considered when the BHEPs are used.

Some Data Related to Use of Displays

The HEPs in this handbook were drawn from all the sources described above as well as from existing tables of error data, when suitable. As an example of the manner in which error probabilities were derived, the following section outlines the derivation of estimates associated with analog and digital displays. BHEPs were derived for the following:

Quantitative reading errors on analog displays

Check-reading errors on analog displays

Quantitative reading errors on digital displays

Recording errors

Data on Quantitative Reading of Analog Displays

In one study of errors in reading and orally reporting two-digit numbers from analog displays (Chapanis and Scarpa, 1967), a response was scored as an error if it was incorrect by a single digit; i.e., if either the first or second digit was incorrect. Under these conditions, the error probability was .05. In another study of errors in reading two- or three-digit numbers from standard dials (Murrell et al, 1958), the subjects were allowed a tolerance of one digit for the least-significant (last) digit. Thus, if the required number was 148 and the subject read 147 or 149, the response was counted as correct. In this study, the error probability was .024. It is obvious that the error probability declines rapidly as the tolerance is increased, which is what we would expect. In a real-world situation, an analog display rarely has to be read to an accuracy of one digit on a scale of 100 or more. The accuracy of the display itself is usually only 3% of the full-scale reading, so that errors of the last digit are of little consequence. Although we do not have experimental data on the frequency of gross errors in reading analog displays, in most cases it is not important that an operator obtain an exact reading, so the error probability reported in laboratory studies can be modified on the basis of industrial requirements.

The latter study cited above was a realistic study of analog display reading. A group of British Naval Chief Petty Officers served as subjects, and read aloud a variety of standard power meters at various distances. The distance of interest to us, 30 inches, is typical of the distance at which operators would read a meter in an NPP. At 30 inches, there were 60 errors in 2520 trials, yielding the HEP of .024. Any error of more than one digit was scored. (About half the readings were two-digit numbers and half were three-digit numbers.)

The report does not indicate the magnitude of errors, but a related study allows us to form an estimate. In a study of visual interpolation (Guttmann and Finley, 1970), it was found that error scores were reduced by more than an order of magnitude when the error tolerance was increased from one-tenth to two-tenths of a scale division. In this study, one-tenth of a scale division corresponded to one digit. Similar findings were obtained in another study (Kappauf and Smith, 1948).

If we apply these findings to the Murrell data and assume an order-of magnitude reduction in error for a one-digit increase in error tolerance, we obtain an HEP of .0024, which is very close to the .003 we used in WASH-1400 and which we have used here. We are using this larger error probability, .003 for reading errors per se, to allow for those cases in which people read meters from distances greater than the 30 inches assumed to be typical.

Data on Check-Reading of Analog Displays

The HEP for check-reading of analog displays has also been studied in laboratory settings. For the conventional single-pointer display with limit marks, the probability of a check-reading error is approximately .001. This estimate is based on the AIR Data Store (Munger et al, 1962). A study by Kurke (1956) indicates that errors occur more than twice as frequently with meters that do not have limit marks on them as compared with meters that are so marked. Therefore, we would expect check-reading errors of analog displays without limit marks to occur with a probability of .002. However, because the check-reading conditions in an NPP are not as well-structured as those in the Kurke study, we assumed an HEP of .003 to allow for various PSFs that could affect reliability.

Data on Quantitative Reading of Digital Displays

For quantitative reading of digital displays, no tolerances are allowed in scoring since no interpolation or alignment of scales and pointers is involved. Digital displays are read more accurately than analog displays. In a study of digital-clock reading (Zeff, 1965), four errors were made in 800 trials, an unmodified error probability of .005. This figure includes errors of reading, memory, and recording. Subjects read the four-digit digital clocks as rapidly as they could, and the display was removed from view as soon as they started writing so there was no opportunity to review the display. In a practical NPP situation a person would have adequate time to read his display and could recheck his readings as desired. Also, each reading would be an individual act, not part of a repetitious series. We estimate that the probability of error on the reading task alone (excluding writing errors) is no greater than .001 in a realistic situation.

As stated in Chapter 11, the operator does not really check-read digital readouts; he must actually read the value. Therefore, whenever he is checking a digital readout, we use the estimated .001 HEP for quantitative reading of this type of display. Inaccuracies in remembering what digital value is appropriate or in reading the desired value from a written procedure are not included in this .001 HEP.

Errors of Recording

The HEP for entering data from a display into a log is actually a composite of at least two actions--reading a display and recording the reading. Sometimes a third action, communicating, is involved, as when the reader states the reading to the writer, but this activity will be disregarded. Thus, there are two primary opportunities for error when a quantitative display indication is to be recorded. In one study, skilled machinists copied 6-digit numbers from their own handwriting onto a special form. Their HEP for reading and recording was .004 (Rigby and Edelman, 1968b).

If we assume the reading HEP to be .001 (see previous page) and subtract this from the above value, the HEP for recording per se is .003. Actually, this value is pessimistic because the errors occurred in the course of performing a series of calculations, and the subjects were not paying particular attention to recording individual groups of numbers accurately. In a realistic NPP situation, in which a person reads a display and records it on a prepared form as an individual task, without time pressure and with opportunity to recheck his readings, errors will occur with an estimated probability no greater than .001. Thus, the combined error probability of reading and recording a digital display (of four digits or less) is estimated as .002.

Some Data Related to Valving Operations

In developing the HEPs for Chapter 13 (Valving Operations), many of them were based on judgments. The BHEP of omission of a step in a procedure (.003, derived from industrial experience, records of inspection, and, field studies in military and industrial settings) was modified (as in Table 13-1, p 13-8), on the basis of applicable PSFs. Similarly, in Table 13-2 (p 13-14) the BHEP of .003 was applied to a "standard" situation: maintainer checking valve status using a short checklist, item within field of view, no checkoff used, and personal safety not affected. The other entries in the table were derived by judging the effects of relevant PSFs.

To estimate the probability that a lock on a valve would not be locked (an error of omission), the BHEP of .003 was modified downward and upward for the situations in which PSFs are conducive to remembering or forgetting the task, respectively.

The estimated probability that plant procedures or policies will be ignored (.01) is based on industrial experience--it reflects the occasional lapse from conscientious performance that occurs in any normal, properly motivated group. Such lapses may occur because of time stress, fatigue, or other causes. The figure does not consider chronic violators of plant policies--it is assumed that chronic violators will not be tolerated.

The estimated .01 probability that plant procedures or policies will be ignored does not apply given the proper use of checklists (Chapters 8 and 14). The use of checklists is a unique case in that people may use a checklist incorrectly without violating plant policies. (This could occur because the plant training program does not stress the proper use of checklists.) Also, the tendency to develop shortcuts in the use of checklists is often regarded as a step toward improved efficiency by the checklist user rather than as a violation of plant policies. The estimate that a checklist will be used properly only one-half of the time is based on observations in a number of industrial settings. (Some reviewers of the handbook have stated that the estimate of correct use of the checklist 50% of the time may be too optimistic, and this estimate is under review.)

Some Data Related to Manual Controls

The HEPs in Chapter 12, "Manual Controls," are based on other data stores and have been modified for NPP situations. The basic situation is that of selecting a control within a group of identical controls identified only by labels. The AIR Data Store (Munger et al, 1962) suggests an error probability of .002 for such a situation. Our HEP of .003 allows for the difficulty of relating controls to their functions in a typical NPP. The error probability is judged to be reduced by a factor of 3 if the controls are functionally grouped, and by another factor of 2 if clearly drawn mimic lines are provided. This is consistent with the improvements expected as PSFs are improved (Chapter 3).

If the desired control is properly selected, the only remaining errors of consequence are the direction of operation of the control and the final control setting. For errors in direction of operation, the estimated HEPs are derived from a review of the available data stores: the Bunker-Ramo Data Bank (Meister, 1967), the Aerojet-General Data Bank (Irwin et al, 1964a and b), and the AIR Data Store (Munger et al, 1962).

For errors in setting a multiposition switch, there are numerous relevant variables that can affect the accurancy of the setting--thus the wide error bounds associated with the HEP. The HEP of .01 for failure to mate a connector properly is based on records of weapons assembly work over a number of years.

A Comment on Errors of Judgment

Assuming that operators have made no perceptual errors, we are unable to estimate errors of judgment (decision-making errors). We classify such errors as sporadic (p 2-16). They occur infrequently, and relatively few cases have been reported. Because of the infrequency of such errors, we do not have enough data to formulate a probability of occurrence. Errors of judgment are difficult to predict if considered apart from related factors such as display adequacy or adequacy of procedures and training. In aircraft accident investigations, errors have often been classified as errors of judgment when they were obviously associated with difficulty in reading a display or interpreting the meaning of a display (Hurst, 1976).

As in the aviation field, errors in NPPs are sometimes classified as errors of judgment although the operators were responding in accordance with inappropriate procedures that had been overemphasized in training. Therefore, the handbook does not offer an HEP for errors of judgment per se. Error bounds associated with actual tasks that could involve errors of judgment are broad enough to include the occasional error of judgment. Until much more data on errors in decision-making are collected from plant and simulator experience, no such HEPs can be formulated.

Comments on Data Sources

The preceding describes the manner in which the HEPs in this handbook were derived. It is not intended to be exhaustive (not all HEPs are discussed); rather, it describes the approach that was followed. We expect that experience will reveal any gross inaccuracies in our estimates, and we fully expect inaccuracies to be discovered. To the extent of our knowledge, we have attempted to err on the conservative side, so our estimates will probably be somewhat pessimistic. In view of the potential consequences of errors, we felt that this was a prudent approach. When sufficient data have been gathered to provide more reliable estimates of the HEPs, we will revise them.

CHAPTER 20. DERIVED HUMAN ERROR PROBABILITIES AND RELATED PERFORMANCE SHAPING FACTORS

This chapter summarizes the estimated HEPs and uncertainty bounds presented in Part III. The tables in this chapter are intended for use as quick references, and are cross-referenced to the chapters from which they are drawn. The user is urged to familiarize himself with the source chapters for the proper use of the error terms and the assumptions on which they are based. The user may need to modify the HEPs if the PSFs for his specific application differ from those assumed in this chapter.

Performance Shaping Factors

All human actions are subject to the effects of PSFs. Chapter 3 describes the usual PSFs that influence HEPs in industrial settings. The HEPs listed are premised on "average" industrial conditions. There is considerable latitude in the word "average," but for practical purposes, average conditions are those which do not subject a worker to an unusual degree of discomfort. For example, the comfort range for temperature is generally specified as 66 to 71°F (19 to 22°C). Although some discomfort is felt outside these limits, most people can function without impairment in the range of 63 to 73°F (17 to 23°C). Therefore, we do not apply any correction factor for the PSF temperature unless the temperature exceeds these limits. Any correction factor should be based on judgment for the particular situation, and should include consideration of the actual temperature, relative humidity, nature of the task, time spent in the situation, etc. For the particular case of temperature, two studies (Osborne and Vernon, 1922; and Vernon and Bedford, 1931) indicate that accident rates increase by about 40% at temperatures over 80°F (27°C).

Not all PSFs can be described as accurately as temperature. For example, as discussed in Chapter 15, "Recovery Factors and Administrative Control," it is difficult to quantify the quality of administrative practices in a plant, or to estimate their overall effect on error probabilities, although these practices do constitute an important PSF. In most cases, the user will have to be content with rating administrative practices as "good," "average," or "poor," making a subjective decision about the effect that this PSF may have on any particular task.

PSFs such as temperature, noise level, and others related to the comfort of the worker will usually be average (or better) in NPPs since regulatory agencies such as OSHA and organizational sectors such as employee unions will promptly report deviations from recommended levels. However, the PSFs related to ergonomics considerations in systems operation are not subject to regulation, and large variations exist from plant to plant as well as within individual plants. The estimated HEPs summarized here are based on the "average" conditions observed in a number of operating plants in the U.S. and Europe -- the error bounds reflect the range of variability in performance attributable to differences in relevant PSFs, differences between (and within) people, and our own uncertainty about the actual HEPs.

How to Use the Data Tables

The method of using estimated HEPs in an analysis is described in detail in Chapters 4 and 5, with further instruction and examples in Chapter 21. The most common error in performing analyses is the failure to consider the effects of dependence between tasks and between people. The user is urged to review Chapter 7 before using the HEPs. Finally, the tables do not stand alone; it is necessary to read the material in each section in this chapter which pertains to the table of interest.

The tables of error probabilities are listed in the following sequence:

Sections	Table	Table Page	From Chapter
Dependence (p 20-4)	20-1, 20-2	20-6 20-7	7 7
Displays (Reading and Recording) (p 20-8)			
Annunciators (p 20-8)	20-3, 20-4	20-9 20-10	10 10
Annunciated Printout Equipment (p 20-8)			
Quantitative Readings (p 20-11)	20-5	20-11	11
Recording Errors (p 20-11)	20-6	20-11	11
Check-Reading Displays for Specific Purpose (p 20-12)	20-7	20-12	11
Inspection Tasks (p 20-13)			
Walk-Around Inspections (p 20-13)	20-8	20-13	8
Checking Other People's Work (p 20-14)	20-9	20-14	15
Initial Audits (p 20-15)	20-10	20-15	9,10,11
Periodic Scanning of Unannunciated Displays (p 20-16)	20-11, 20-12	20-17, 20-18	9,11 9,11
Manual Operation of Controls (p 20-19)	20-13	20-19	12
Valves (p 20-20)			
Errors of Commission by Operator (p 20-21)	20-14	20-21, 20-22	13
Errors of Omission by Operator (p 20-23)	20-15	20-23	13
Errors of Recovery (p 20-24)	20-16	20-25	13

Sections	Table	Table Page	From Chapter
Errors of Maintainers (p 20-24)	20-17	20-26	13
Errors of Supervisors (p 20-24)			
Procedures (p 20-27)			
Oral Instructions (p 20-27)	20-18	20-28	14
Preparation of Written Procedures (p 20-28)	20-19	20-28	14
Use of Written Procedures (p 20-29)	20-20	20-29	7,14,15
Administrative Control (p 20-30)	20-21, 20-22	20-30 20-31	15 15
Stress and Skill Levels (p 20-31)	20-23, 20-24	20-32, 20-33	17,18 17
HEPs from WASH-1400 (p 20-33)	20-25	20-34	
Uncertainty Bounds (p 20-35)	20-26	20-36	16
Graphic Representation of HEPs (p 20-37)	Fig. 20-1 20-27	20-38, 20-39, 20-40, 20-41	

From this point on we will omit the word "estimated," but it should be understood that all listed HEPs are estimated HEPs. Confidence in these HEPs is reflected in the width of the uncertainty bounds associated with each term. As an aid to finding specific sections which provide the rationales for the HEPs, the appropriate page numbers are cited with the HEPs in the text or in the tables. The HEPs apply to the case of an individual operator. The modifications of HEPs when two or more operators are available are described in Chapters 14 and 15.

Dependence (Chapter 7)

Analysts are urged to estimate conditional probabilities on the basis of data from the work situation. When such data are not available, the dependence model can be used to approximate the conditional probabilities of task success or failure. In this model, all of the conditioning influence on a task is assumed to originate only from the performance of the immediately preceding task, or from some influence common to both tasks. The continuum of dependence between two tasks is represented by five points: zero dependence (ZD), low dependence (LD), moderate dependence (MD), high dependence (HD), and complete dependence (CD). For ZD, the conditional probability of the second task is the basic probability of success or failure. For CD, the conditional probabilities are always 1.0. For the three intermediate levels of dependence (low, moderate, and high), the following equations apply (where BHSP is the basic human success probability and BHEP is the basic human error probability):

$\Pr\{S\|S \text{ on the previous task}\}$	Level of Dependence	$\Pr\{F\|F \text{ on previous task}\}$
$\frac{1 + 19(\text{BHSP})}{20}$	LD	$\frac{1 + 19(\text{BHEP})}{20}$
$\frac{1 + 6(\text{BHSP})}{7}$	MD	$\frac{1 + 6(\text{BHEP})}{7}$
$\frac{1 + \text{BHSP}}{2}$	HD	$\frac{1 + \text{BHEP}}{2}$

For BHEPs of .01 or smaller, the conditional HEPs are approximated by the values of .05, .15, and .5 for the conditions of LD, MD, and HD, respectively.

Tables 20-1 and 20-2 list the conditional probabilities of success and failure for a range of HEPs, under the five levels of dependence.

Table 20-1. Conditional Probabilities of Success or Failure for Task "N" for the Five Levels of Dependence, Given Failure on Task "N-1" (from Table 7-2)

Task "N" Conditional Probabilities*

ZD**		LD		MD		HD		CD	
S	F	S	F	S	F	S	F	S	F
.75	.25	.71	.29	.64	.36	.37	.63	0	1.0
.9	.1	.85	.15	.77	.23	.45	.55	0	1.0
.95	.05	.9	.1	.81	.19	.47	.53	0	1.0
.99	.01	.94	.06	.85	.15	.49	.51	0	1.0
.995	.005	.95	.05	.85	.15	.50	.50	0	1.0
.999	.001	.95	.05	.86	.14	.50	.50	0	1.0
.9995	.0005	.95	.05	.86	.14	.50	.50	0	1.0
.9999	.0001	.95	.05	.86	.14	.50	.50	0	1.0
.99999	.00001	.95	.05	.86	.14	.50	.50	0	1.0

*All probabilities are rounded. Equations 7-14 through 7-18 were used to calculate the values in the F columns. The values in the S columns were obtained by subtraction.

**The conditional probabilities given ZD are the basic probabilities for Task "N."

Table 20-2. Conditional Probabilities of Success or Failure for Task "N" for the Five Levels of Dependence, Given Success on Task "N-1" (from Table 7-3)

Task "N" Conditional Probabilities*

ZD**		LD		MD		HD		CD	
S	F	S	F	S	F	S	F	S	F
.75	.25	.76	.24	.79	.21	.87	.13	1.0	0
.9	.1	.9	.1	.91	.09	.95	.05	1.0	0
.95	.05	.95	.05	.94	.06	.97	.03	1.0	0
.99	.01	.99	.01	.991	.009	.995	.005	1.0	0
.995	.005	.995	.005	.996	.004	.997	.003	1.0	0
.999	.001	.999	.001	.999	.001	.9995	.0005	1.0	0
.9995	.0005	.9995	.0005	.9996	.0004	.9997	.0003	1.0	0
.9999	.0001	.9999	.0001	.99991	.00009	.99995	.00005	1.0	0
.99999	.00001	.99999	.00001	.999991	.000009	.999995	.000005	1.0	0

*All conditional probabilities are rounded. Equations 7-9 through 7-13 were used to calculate the values in the S columns. The values in the F columns were obtained by subtraction.

**The conditional probabilities given ZD are the basic probabilities for Task "N."

Displays (Reading and Recording)

Annunciators (Chapter 10)

The term annunciated displays includes all visual indicators that are announced by a compelling auditory signal. Most of the annunciated displays in an NPP are legend lights, but there also are printouts and other indicators.

In Tables 20-3 and 20-4, the term one annunciator includes a functional grouping of annunciators consisting of more than one annunciator with complete dependence among them. As the number of annunciators (or functional groups of annunciators) sounding in a very brief interval increases, the probability of failure to respond to any one of them increases as shown in Table 20-4.

During the performance of certain calibration or maintenance procedures an alarm will sound a number of times as adjustments are made. These alarms, although anticipated by the operator, constitute "false alarms," and the operator may turn off the audio and visual signals without verifying the legend light. In such cases, the probability of failure to respond to an actual alarm is .001 (.0001 to .01) (p 10-22).

Annunciated Printout Equipment (Chapter 10)

The HEPs for failure to respond to annunciated printout equipment are the same as for annunciated legend lights. The reading error for the general sense of the message is negligible, as there is no confusion about which message to read. It is judged that if the message is read at all, it will be read correctly. This excludes errors made in reading coded messages or series of numbers for which quantitative reading errors are possible (see next section).

Table 20-3. Probabilities of Error for Annunciated Legend Lights

Task	HEP
Respond to an annunciated legend light (one of one) (p 10-9)	.0001 (.00005 to .001)
Reading the message (this figure includes the probability of reading the wrong legend light) (p 10-6)	.001 (.0005 to .005)
Resume attention to a legend light within 1 minute after an interruption (sound and blinking cancelled before interruption) (p 10-11)	.001 (.0001 to .01)
Respond to a legend light if more than 1 minute elapses after an interruption (sound and blinking cancelled before interruption) (p 10-11)	.95 (.9 to .99)

Table 20-4. Probability of Failure to Respond to One Randomly Selected Annunciator of Several (from Table 10-2, right column)

Number of Annunciators	HEP
1	.0001 (.00005 to .001)
2	.0006 (.00006 to .006)
3	.001 (.0001 to .01)
4	.002 (.0002 to .02)
5	.003 (.0003 to .03)
6	.005 (.0005 to .05)
7	.009 (.0009 to .09)
8	.02 (.002 to .2)
9	.03 (.003 to .3)
10	.05 (.005 to .5)
11-15	.10 (.01 to .999)
16-20	.15 (.015 to .999)
21-40	.20 (.02 to .999)
> 40	.25 (.025 to .999)

Quantitative Readings (Chapter 11)

Table 20-5. Probabilities of Errors of Commission in Reading Quantitative Information from Displays

Reading Task	HEP
Analog meter (p 11-7)	.003 (.001 to .01)
Digital readout (p 11-7)	.001 (.0005 to .005)
Chart recorder (p 11-7)	.006 (.002 to .02)
Printing recorder with large number of parameters (p 11-8)	.05 (.01 to .2)
Graphs (p 11-8)	.01 (.005 to .05)
Values from indicator lamps that are used as quantitative displays (p 11-8)	.001 (.0005 to .005)
Recognize that an instrument being read is jammed, if there are no indicators to alert the user (p 11-11)	.1 (.02 to .2)

Recording Errors (Chapter 11)

The probability that an entry will be recorded incorrectly on some data form is obtained by adding the reading HEPs from Table 20-5 to the recording HEPs listed in Table 20-6.

Table 20-6. Probabilities of Errors of Commission in Recording Readings (p 11-7)

Number of Digits to be Recorded	HEP
≤ 3	Negligible
> 3	.001 (.0005 to .005)

Check-Reading Displays for Specific Purpose (Chapter 11)

Table 20-7. Probabilities of Errors of Commission in Check-Reading Displays

Check-Reading Task	HEP
Digital indicators (these must be read -- there is no true check-reading function for digital displays) (p 11-14)	.001 (.0005 to .005)
Analog meters with easily seen limit marks (p 11-12)	.001 (.0005 to .005)
Analog meters with difficult-to-see limit marks, such as scribe lines (p 11-14)	.002 (.001 to .01)
Analog meters without limit marks (p 11-12)	.003 (.001 to .01)
Analog-type chart recorders with limits (p 11-12)	.002 (.001 to .01)
Analog-type chart recorders without limit marks (p 11-12)	.006 (.002 to .02)
Confirming a status change on a status lamp (p 11-15)	Negligible
Checking the wrong indicator lamp (in an array of lamps) (p 11-15)	.003 (.001 to .01)
Misinterpreting the indication on the indicator lamps (p 11-15)	.001 (.0005 to .005)

Inspection Tasks

The term inspection includes all monitoring activities for detecting conditions that are out of limits or approaching a limit. Inspection includes reading displays as well as all forms of verification such as the status of switches, valves, and indicators; and general observation of plant status, as in the performance of a walk-around inspection. It is assumed that if a deviant item is not detected in 30 days, it will not be detected until some other attention-getting event occurs.

Walk-Around Inspections (Chapter 8)

Table 20-8 indicates the increase in the HEP for the case of a single operator performing one walk-around daily without a checklist. The HEPs apply to the detection of an incorrect state of an item such as a manual valve, if there are no indicators of the correct state. For the case in which more than one operator performs a walk-around and for cumulative probabilities of detection, see Chapter 8.

Table 20-8. Probabilities of Error in Walk-arounds,
One Operator, No Checklist (from Table 8-1)

Day #	HEP
1	.9 (.5 to .99)
2	.95 (.6 to .995)
3	.975 (.7 to .999)
4	.99 (.8 to .999)
5-30	.999 (.9 to .999)
> 30	1.0

Checking Other People's Work (Chapter 15)

In addition to the walk-around, the performance of operators and others is often checked by people we call checkers or inspectors. The following table presents HEPs for some checking tasks. (See also Tables 20-21 and 20-22.)

Table 20-9. Probabilities that a Checker will Fail to Detect Errors (from Table 15-1)

Checking Operation	HEP
Usual monitoring in NPP with some kind of checklist or written procedure (includes tasks such as over-the-shoulder checking and checking written lists or procedures)	.10 (.05 to .5)
Same as above but without written materials	.20 (.10 to .9)
Special short-term, one-of-a-kind checking (e.g., supervisor checks performance of a novice)	.05 (.01 to .10)
Hands-on type of checking that involves special measurements or other activities	.01 (.005 to .05)
Repeated checking of one person's work by different individuals during or after completion of a standard or routine task	Assume HD among checkers; no recovery credit for more than 2 checkers
Two-man team, with one person the doer and the other the reader/checker	Assume HD between doer and reader/checker
Checking of valve status	Use HEPs from Tables 20-16 and 20-17
Walk-around checking operation	Use HEPs from Table 20-8

Initial Audits (Chapters 9, 10, and 11)

The initial audit is made by the oncoming shift when taking over from the previous shift. The HEPs below apply to the case of a single operator conducting an audit. The cases of more than one operator are discussed in Chapters 9, 14, and 15. In the sections below on detection of deviant indications, the following assumptions are made:

1. Detection of a deviant indication is regarded as a simple task, with equal probabilities of detection for all deviant indications of the same type of display.
2. There is no observer uncertainty associated with the indications which direct attention to certain displays. (There is no ambiguity as to the response required.)
3. There are no alarms operating to direct attention to certain displays.
4. A deviation not detected in 30 days will not be detected in later periodic scanning.

Table 20-10. Probabilities of Error on Initial Audit
No Special Alerting Cues and Only One Deviant Display

Task	HEP
Deviant meter with limit marks (p 11-21)	.05 (.01 to .1)
Deviant meter without limit marks (p 11-26)	.15 (.03 to .3)
Annunciator light requiring action, which is no longer annunciating (p 10-7)	.9 (.8 to .98)
Incorrect status of a legend light, other than an annunciator light (p 11-17)	.98 (.96 to .996)
Incorrect status of an indicator lamp (p 11-17)	.99 (.98 to .998)

For the case of more than one deviant indication, the probability of detecting at least one is a function of the HEP for a single deviant indication (Equation 9-1, p 9-13). Some representative probabilities are listed in the next section.

Periodic Scanning of Unannunciated Displays (Chapters 9 and 11)

The previous section lists the HEPs for the initial audit. If additional scans are made, the probability of failure to detect a deviant display in any given scan is listed in Table 20-11; eight hourly scans are assumed beginning with the initial audit. Again, the 30-day cutoff must be used in applying the HEPs.

Table 20-11. Probabilities of Failure to Detect at Least One of One to Five Deviant Unannunciated Displays (Condensed from Table 9-1)

Number of Deviant Indications

1	2	3	4	5
.99	.985	.98	.975	.97
.95	.93	.90	.88	.86
.90	.85	.81	.77	.73
.80	.72	.65	.58	.52
.70	.59	.51	.43	.37
.60	.48	.39	.31	.25
.50	.37	.28	.21	.16
.40	.28	.20	.14	.10
.30	.19	.13	.08	.05
.20	.12	.07	.04	.03
.10	.05	.03	.02	.01
.05	.03	.01	.007	.004
.01	.005	.003	.001	.001

Uncertainty Bounds

For HEPs $< .5$,

Lower Bound = HEP x .2

Upper Bound = HEP x 2

For HEPs $\geq .5$

Lower Bound = 1 - 2(1-HEP)

Upper Bound = 1 - .2(1-HEP)

Table 20-12. Probabilities of Failure to Detect One (of One) Deviant Unannunciated Display* at Each Scan, When Scanned Hourly**

	Hourly Scans							
Display Type	1	2	3	4	5	6	7	8
Meter with limit marks (p 11-21)	.05	.31	.50	.64	.74	.81	.86	.90***
Meter without limit marks (p 11-26)	.15	.47	.67	.80	.87	.92	.95	.97
Chart recorders with limit marks (p 11-26)	.10	.40	.61	.74	.83	.89	.92	.95
Chart recorders without limit marks (p 11-26)	.30	.58	.75	.85	.91	.94	.97	.98
Annunciator light no longer annunciating (p 10-7)	.9	.95	.95	.95	.95	.95	.95	.95
Legend light other than annunciator light (p 11-17)	.98	.98	.98	.98	.98	.98	.98	.98
Indicator lamp (p 11-17)	.99	.99	.99	.99	.99	.99	.99	.99
Digital readout (p 11-18)	(These are not part of the scanning model, because they must be read. See Table 20-5.)							

---

*One display refers to a single display or a group of completely dependent displays.

**For fewer than 8 hourly scans, see the special instructions on p 9-6.

***In estimating uncertainty bounds for HEPs $< .5$, lower bound = HEP x 0.2 and upper bound = HEP x 2. For HEPs $\geq .5$, lower bound = 1 - 2(1-HEP) and upper bound = 1 - 0.2(1-HEP).

Manual Operation of Controls (Chapter 12)

Manual operation of controls includes the operation of all kinds of switches, connectors, and valves. Table 20-13 applies to controls other than valves and lists errors of commission only. For errors of omission, use the applicable HEPs in the subsequent tables on valves. If controls are handled as pairs, assume CD between them. The effects of tagging are described in the following section, "Valves."

Table 20-13. Probabilities of Errors of Commission in Operating Manual Controls (from Table 12-1)

Task	HEP
Select wrong control in a group of identical controls identified by labels only	.003 (.001 to .01)
Select wrong controls from a functionally grouped set of controls	.001 (.0005 to .005)
Select wrong control from a panel with clearly drawn mimic lines	.0005 (.0001 to .001)
Turn control in wrong direction (no violation of populational stereotypes)	.0005 (.0001 to .001)
Turn control in wrong direction under normal operating conditions (violation of a strong populational stereotype)	.05 (.01 to .1)
Turn control in wrong direction under high stress (violation of a strong populational stereotype)	.5 (.1 to .9)
Set a multiposition selector switch to an incorrect setting	.001 (.0001 to .1)*
Improperly mate a connector	.01 (.005 to .05)

---

*The unusually wide error bounds reflect the wide variety of designs, ranging from good to barely acceptable ergonomics.

Valves (Chapter 13)

This section presents the HEPs related to operation of two types of valves: (1) a locally-operated valve that is opened or closed by manipulating a handle such as a wheel and (2) a motor-operated valve that is remotely controlled by a switch in the control room. We designate the first type as a manual valve and the second type as an MOV. Where appropriate, the HEPs in Table 20-13 (Manual Controls) can be applied to valves.

It is assumed that the changing of valves from their normal operating positions and their subsequent restoration is normally performed by operators. Operations performed by maintenance personnel within some subsystem previously valved off by operating personnel are exceptions.

There are three levels of tagging and locking (Table 15-3, p 15-11). Unless otherwise stated, the HEPs in the following tables are premised on the use of a Level 2 tagging or locking system. With Level 2 and Level 3 tagging, ZD is assumed between task steps. With Level 1 tagging, CD is assumed between all items for which tags are prepared, for errors of omission. If no tags are used, multiply the HEPs associated with Level 2 tagging by a factor of 10, with a maximum HEP of .5.

When both tagging and locking systems are used, one identifies the highest level of control and bases the analysis on that level. No extra credit is allowed for both, even when both are Level 1.

The HEPs for checking tasks related to valves include the effects of dependence, and can be used without modification.

Errors of Commission by Operator

Table 20-14. Probabilities of Errors of Commission by Operator Changing or Restoring Valves

Task	HEP
Writing any one item when preparing a list of valves (or tags) (p 13-5)	.003 (.001 to .01)
Change or tag wrong valve where the desired valve is one of two or more adjacent, similar-appearing manual valves, and at least one other valve is in the same state as the desired valve, or the valves are MOVs of such type that valve status cannot be determined at the valve itself (p 13-5)	.005 (.002 to .02)
Restore the wrong manual valve where the desired valve is one of two or more adjacent, similar-appearing valves, and at least two are tagged out for maintenance (p 13-5)	.005 (.002 to .02)
Reversal error: change a valve, switch, or circuit breaker that has already been changed and tagged (p 13-5)	.0001 (.00005 to .001)
Reversal error, as above, if the valve has been changed and NOT tagged (p 13-5)	.1 (.01 to .5)
Note that there is more than one tag on a valve, switch, or circuit breaker that he has decided to restore (p 13-5)	.0001 (.00005 to .0005)
Change or restore wrong MOV switch or circuit breaker in a group of similar-appearing items (In case of restoration, at least two items are tagged) (p 13-6)	.003 (.001 to .01)

(Table concluded on next page.)

Task	HEP
Complete a change of state of an MOV of the type that requires the operator to hold the switch until the change is completed as indicated by lamp (p 13-6)	.003 (.001 to .01)
Given that a manual valve sticks, operator erroneously concludes that the valve is fully open (or closed) (p 13-6):	
Rising-stem valves	
If the valve sticks at about three-fourths or more of its full travel (no position indicator* present)	.005 (.002 to .02)
If there is an indicator showing the full extent of travel	.001 (.0005 to .01)
All other valves	
If there is a position indicator on the valve	.001 (.0005 to .01)
If there is a position indicator located elsewhere (and extra effort is required to look at it)	.002 (.001 to .01)
If there is no position indicator	.01 (.003 to .1)

---

*A position indicator is a scale that indicates the position of the valve relative to a fully opened or fully closed position.

Errors of Omission by Operator

Table 20-15. Probabilities of Errors of Omission by Operators Changing or Restoring Valves

Task	HEP
Omit an item when preparing a list of valves or set of tags (p 13-7)	.003 (.001 to .01)
Failure to carry out a specific oral instruction to change or restore a valve (for more than one oral instruction use Table 20-18)	.001 (.0005 to .005)
With Level 1 tagging, CD is assumed between all steps in a task for errors of omission -- if the first step is performed, all steps will be performed. The only error of omission is the failure to initiate the task (p 13-7)	.001 (.0005 to .005)
Omit a particular valve change or restoration:	
using written procedures	Table 20-20
not using written procedures	Table 20-18
Checkoff provision of procedure is misused (check off several items at a time) (p 13-7)	.5 (.1 to .9)
Failure to follow established procedures or policies in valve changes or restoration (p 13-7)	.01 (.005 to .05)
Failure to tag a valve, circuit breaker, or switch after changing it (p 13-9)	.001 (.0005 to .05)
Failure to lock a valve after restoration (if required) (p 13-9)	.003 (.001 to .01)
Failure to lock a valve after restoration (if required), with Level 1 lock and key control (p 13-9)	.001 (.0005 to .005)

Errors of Recovery

The HEPs in Table 20-16 below are premised on Level 2 tagging and locking systems -- do not use for Level 1 tagging. Assume ZD between all items in a task except those treated as a unit.

When basic HEPs are to be divided or multiplied by some factor, so are the bounds (with an upper limit of .999).

Errors of Maintainers

Table 20-17 includes errors by the maintainer not included in Table 20-16. The motivation of maintainers is directed primarily at performing maintenance (p 13-12). Restoration of valves after maintenance is not part of this primary concern. The estimated probability for a maintainer of failing to restore an untagged valve after maintenance is .05 (.01 to .1) (p 13-15). The motivation of the maintainer to ensure that a system is correctly valved out to permit maintenance is considered to be stronger than that for restoration, and the HEPs are presented in the table below.

Errors of Supervisors

An HEP of .9 (.80 to .99) is assigned to the case in which a supervisor believes a specific oral instruction was carried out when it was not (p 13-15). Normally, this factor is not included in a human reliability analysis.

Table 20-16. Probabilities of Errors by Operators or Maintainers Checking Valve Operations Performed by Others

Task	HEP
Failure to detect an error of commission (p 13-9)	Use the basic HEP times 10 (upper limit of .9)
Failure to detect an error of omission (p 13-10)	.1 (.05 to .5)
Failure to detect errors of commission or omission on a list of valves or set of tags (p 13-10)	.1 (.05 to .5)
Failure to detect that a manual valve was not completely "home" after being changed (p 13-11)	.5 (.1 to .9)
Failure to detect that a manual valve was not completely "home" after being restored (p 13-11):	
Valves with position indicators	Basic HEP x 10 (upper limit of .9)
Valves without position indicators	Basic HEP x 20 (upper limit of .9)
Failure of a maintainer to detect a reversal error by an operator changing valves (p 13-12)	.5 (.1 to .9)
For the case of an operator checking a maintainer, divide the above HEPs by two (p 13-12)	
If recovery procedures consist of checking paper work only, allow no credit for recovery (p 13-12)	

Table 20-17. Probabilities that Maintainer Will Fail to Check Valve Status Before Maintenance* (from Table 13-2)

Type of List Used

Factor	Short List**			NPP Procedures or Long List**			No List+	
Checkoff Used	Yes++	No		Yes++	No			
Within Field of View		Yes	No		Yes	No	Yes	No
Personal Safety Affected								
Yes	.0004	.0006	.0008	.0005	.0006	.0008	.0008	.001
No	.001	.003	.005	.002	.005	.01	.01	.01

NOTES: *In view of speculative nature of the HEPs, estimates of uncertainty bounds are even more speculative. For HEPs greater than 10^{-3}, use a divisor of 5 for the lower bound and a factor of 5 for the upper bound. For HEPs of 10^{-3} or smaller, use a factor of 5 for the upper bound and a constant 10^{-4} for the lower bound.

**Assume ZD between valves.

+For more than one valve, Table 20-18, indicates the HEPs increase as the number of items to remember increases.

++Credit for checkoff must be modified by the percentage of people who fail to use checkoff properly (Table 20-20).

Procedures (Chapter 14)

This section lists the HEPs associated with various types of procedures including oral, written, and administrative procedures. Many of these HEPs apply to the tasks described in other sections in this chapter.

Oral Instructions

Oral instructions are short instructions given by someone in authority to an operator or maintainer. An oral instruction may include one or more special items. An example of an oral instruction is, "Restore the valves for System ABC." It is up to the recipient of the instruction to find out what the specific valves are. If the person in authority specifies each valve to be restored, each valve is considered to be a separate special instruction item. The probability of forgetting to initiate the task is the probability of forgetting one (of only one) special instruction item.

Table 20-18. Probabilities of Errors in Recalling Special Instruction Items Given Orally (from Table 14-1)

Task	HEP
Items not Written Down by Recipient	
Recall any given item, given the following number of items to remember:	
1 (same as failure to initiate task)	.001 (.0005 to .005)
2	.003 (.001 to .01)
3	.01 (.005 to .05)
4	.03 (.01 to .1)
5	.1 (.05 to .5)
Recall any item if supervisor checks to see that the task was done	Negligible
Items Written Down by Recipient	
Recall any item (exclusive of errors in writing - see Table 20-19)	.001 (.0005 to .005)

Preparation of Written Procedures

In this context, written procedures include any written materials (e.g., tags filled in prior to valve change). The errors of providing incomplete or misleading information are not addressed in this handbook.

Table 20-19. Probabilities of Error in Preparation of Written Procedures

Task	HEP
Omitting an item (p 14-7)	.003 (.001 to .01)
Writing an item incorrectly (p 14-7)	.003 (.001 to .01)

Use of Written Procedures (Chapters 7, 14, and 15)

Table 20-20. Probabilities of Errors of Omission in Use of Written Procedures in Nonpassive Tasks (from Tables 14-2, 14-3, and 15-2)

Task	HEP
Procedures with checkoff provisions* (p 14-11)	
Short list ≤ 10 items	.001 (.0005 to .005)
Long list > 10 items	.003 (.001 to .01)
Checkoff provision improperly used (p 15-9) (Consider procedures with improperly used checkoff provisions to be the same as procedures with no checkoff provisions.)	.5 (.1 to .9)
Procedures with no checkoff provisions* (p 14-11)	
Short list ≤ 10 items	.003 (.001 to .01)
Long list > 10 items	.01 (.005 to .05)
Performance of simple arithmetic calculations (p 14-15)	.01 (.005 to .05)
If two people use written procedures correctly - one reading and checking the other doing the work - assume HD between their performances (p 15-3)	
Procedures available but not used (p 15-10)	
Maintenance tasks	.3 (.05 to .9)
Valve change or restoration tasks	.01 (.005 to .05)

---

*Assume ZD between written steps (p 14-9)

Administrative Control (Chapter 15)

In Table 20-21 there is an entry for failure to use a checklist properly. Such a failure can be ascribed to poor administrative control. Table 20-22 lists the HEPs related to failure of administrative control. These failures can have sizable effects, and must not be ignored in a reliability analysis. See Table 20-9 for errors made by a checker and p 20-20 for HEPs related to type of tagging system employed.

Table 20-21. Estimated Probabilities of Error in Use of Checklists for Passive Tasks Such as the Walk-Around Inspection (from Table 14-2 and Chapter 15)

Task	HEP
Use checklist properly (p 15-9)	.5 (.1 to .9)
Recognize an incorrect status when using a checklist properly (p 14-11)	.01 (.005 to .05)
Recognize an incorrect status when using a checklist improperly (p 14-11)	.1 (.05 to .5)
If two people use checklist properly, one reading, the other inspecting, assume HD (p 15-3)	

Table 20-22. HEPs Related to Failure of Administrative Control (from Table 15-2)

Operation	HEP
Carry out a plant policy when there is no check on a person	.01 (.005 to .05)
Initiate a checking function	.001 (.0005 to .005)
Use control room written procedures under operating conditions that are	
Normal	.01 (.005 to .05)
Abnormal	no basis for estimate
Use a valve restoration list	.01 (.005 to .05)
Use written calibration procedures	.05 (.01 to .1)
Use written maintenance procedures when available	.3 (.05 to .9)
Use a checklist properly (i.e., perform one step and check it off before proceeding to next step)	.5 (.1 to .9)*

---

*This estimate is considered by some to be optimistic.

Stress and Skill Levels (Chapters 17 and 18)

Estimates of HEPs for other than optimal stress levels are highly speculative. Table 20-23 presents some values to use for bounding studies. Table 20-24 presents some assumptions about the effects of multiple operators in a control room following an abnormal event such as a transient.

Table 20-23. Probabilities of Error and Uncertainty Bounds for Levels of Stress for Experienced Personnel and Novices (from Chapter 17 and Table 18-1)

Stress Level	HEPs*	Uncertainty Bounds
	Experienced Personnel (Ch. 17)	
Very Low (p 17-4)	2 x tabled HEPs**	2 x tabled values**
Optimum (p 17-6)	Tabled HEPs	Tabled values
Moderately High (p 17-7):		
Step-by-step tasks	2 x tabled HEPs	2 x tabled values
Dynamic tasks	5 x tabled HEPs	5 x tabled values
Extremely High (p 17-7)	.25	.03 to .75
	Novices*** (Ch. 18)	
Very Low	2 x tabled HEPs**	2 x tabled values**
Optimum		
Step-by-step tasks	Tabled HEPs	Tabled values
Dynamic tasks	2 x tabled HEPs	2 x tabled values
Moderately High		
Step-by-step tasks	4 x tabled HEPs	4 x tabled values
Dynamic tasks	10 x tabled HEPs	10 x tabled values
Extremely High	.25	.03 to .75

*Absolute lowest HEP is 5×10^{-5} and absolute highest HEP is .999.

**Tabled HEPs and uncertainty bounds refer to those in the other tables in Chapter 20.

***Novices are defined as NPP personnel with less than 6 months experience in the jobs for which they are qualified.

Table 20-24. Effects of Several Operators in Control Room in an Abnormal Situation (p 17-24)

Time After Initiation of Abnormal Event	Number of Operators Present and Level of Dependence between Them
0 - 5 min	Only on-duty control room reactor operator is present.
5 - 20 min	Shift supervisor is now present. Assume MD-to-HD between regular operator and shift supervisor. If regular operator is a novice, assume shift supervisor takes over.
20 min - 2 hr	A third reactor operator is now present. Assume HD-to-CD between him and the other two operators.
After 2 hr	Additional qualified personnel are present. We cannot assess their influence; it may be positive or negative.

HEPs from WASH-1400

Table 20-25 is from a copy of Table III 6-1 from Appendix III to WASH-1400. These estimates are still regarded as valid, and are reproduced for use in cases not covered elsewhere. (Note that this table uses the term error rate rather than the equivalent term error probability.)

Table 20-25. General Error Rate Estimates from Table III 6-1 in WASH-1400(a,b)

Estimated Rates	Activity
10^{-4}	Selection of a key-operated switch rather than a non-key switch (this value does not include the error of decision where the operator misinterprets situation and believes key switch is correct choice).
10^{-3}	Selection of a switch (or pair of switches) dissimilar in shape or location to the desired switch (or pair of switches), assuming no decision error. For example, operator actuates large handled switch rather than small switch.
3×10^{-3}	General human error of commission, e.g., misreading label and therefore selecting wrong switch.
10^{-2}	General human error of omission where there is no display in the control room of the status of the item omitted, e.g, failure to return manually operated test valve to proper configuration after maintenance.
3×10^{-3}	Errors of omission, where the items being omitted are embedded in a procedure rather than at the end as above.
3×10^{-2}	Simple arithmetic errors with self-checking but without repeating the calculation by re-doing it on another piece of paper.
$1/x$	Given that an operator is reaching for an incorrect switch (or pair of switches), he selects a particular similar appearing switch (or pair of switches), where x = the number of incorrect switches (or pair of switches adjacent to the desired switch (or pair of switches). The $1/x$ applies up to 5 or 6 items. After that point the error rate would be lower because the operator would take more time to search. With up to 5 or 6 items he doesn't expect to be wrong and therefore is more likely to do less deliberate searching.
10^{-1}	Given that an operator is reaching for a wrong motor operated valve MOV switch (or pair of switches), he fails to note from the indicator lamps that the MOV(s) is (are) already in the desired state and merely changes the status of the MOV(s) without recognizing he has selected the wrong switch(es).
~ 1.0	Same as above, except that the state(s) of the incorrect switch(es) is (are) not the desired state.
~ 1.0	If an operator fails to operate correctly one of two closely coupled valves or switches in a procedural step, he also fails to correctly operate the other valve.

(Table concluded on next page.)

Table 20-25. General Error Rate Estimates from Table III 6-1 in WASH-1400[(a,b)]

Estimated Rates	Activity
10^{-1}	Monitor or inspector fails to recognize initial error by operator. Note: With continuing feedback of the error on the annunciator panel, this high error rate would not apply.
10^{-1}	Personnel on different work shift fail to check condition of hardware unless required by check list or written directive.
5×10^{-1}	Monitor fails to detect undesired position of valves, etc., during general walk-around inspections, assuming no check list is used.
.2 - .3	General error rate given very high stress levels where dangerous activities are occurring rapidly.
$2^{(n-1)}x$	Given severe time stress, as in trying to compensate for an error made in an emergency situation, the initial error rate, x, for an activity doubles for each attempt, n, after a previous incorrect attempt, until the limiting condition of an error rate of 1.0 is reached or until time runs out. This limiting condition corresponds to an individual's becoming completely disorganized or ineffective.
~1.0	Operator fails to act correctly in the first 60 seconds after the onset of an extremely high stress condition, e.g., a large LOCA.
9×10^{-1}	Operator fails to act correctly after the first 5 minutes after the onset of an extremely high stress condition.
10^{-1}	Operator fails to act correctly after the first 30 minutes in an extreme stress condition.
10^{-2}	Operator fails to act correctly after the first several hours in a high stress condition.
x	After 7 days after a large LOCA, there is a complete recovery to the normal error rate, x, for any task.

(a) Modification of these underlying (basic) probabilities were made on the basis of individual factors pertaining to the tasks evaluated.

(b) Unless otherwise indicated, estimates of error rates assume no undue time pressures or stresses related to accidents.

Uncertainty Bounds (Chapter 16)

The tables in this chapter include both HEPs and uncertainty bounds. Table 20-26 is included for those cases in which uncertainty bounds must be determined for tasks not included in the handbook.

Table 20-26. General Guidelines for Estimating Uncertainty Bounds for Estimated HEPs* (from Table 16-2)

Task and/or HEP Guidelines	Uncertainty Bounds**	
	Lower	Upper
HEP-Oriented		
Estimated HEP < .001	HEP ÷ 10	HEP x 10
Estimated HEP .001 to .01	HEP ± 3	HEP x 3
Estimated HEP > .01	HEP ÷ 5	HEP x 2 to x 5
Task-Oriented		
Task consists of step-by-step procedure conducted under routine circumstances and is essentially static (e.g., calibration task), HEP $\geq$ .001.	HEP ÷ 3 to ÷ 5	HEP x 3 to x 5
Same as above, but estimated HEP is < .001.	HEP ÷ 10	HEP x 10
Task consists of step-by-step procedure, but is carried out in nonroutine circumstances specifically involving a potential turbine/reactor trip.	HEP ÷ 5	HEP x 10
Task consists of relatively dynamic interplay between operator and system indications, but task is done under routine conditions; e.g., increasing or reducing power.	HEP ÷ 10	HEP x 10
Task is performed under severe stress conditions; e.g., large LOCA, conditions in which the status of ESFs is not perfectly clear, or conditions in which the initial operator responses have proved to be inadequate and now severe time pressure is felt.	.03	.75

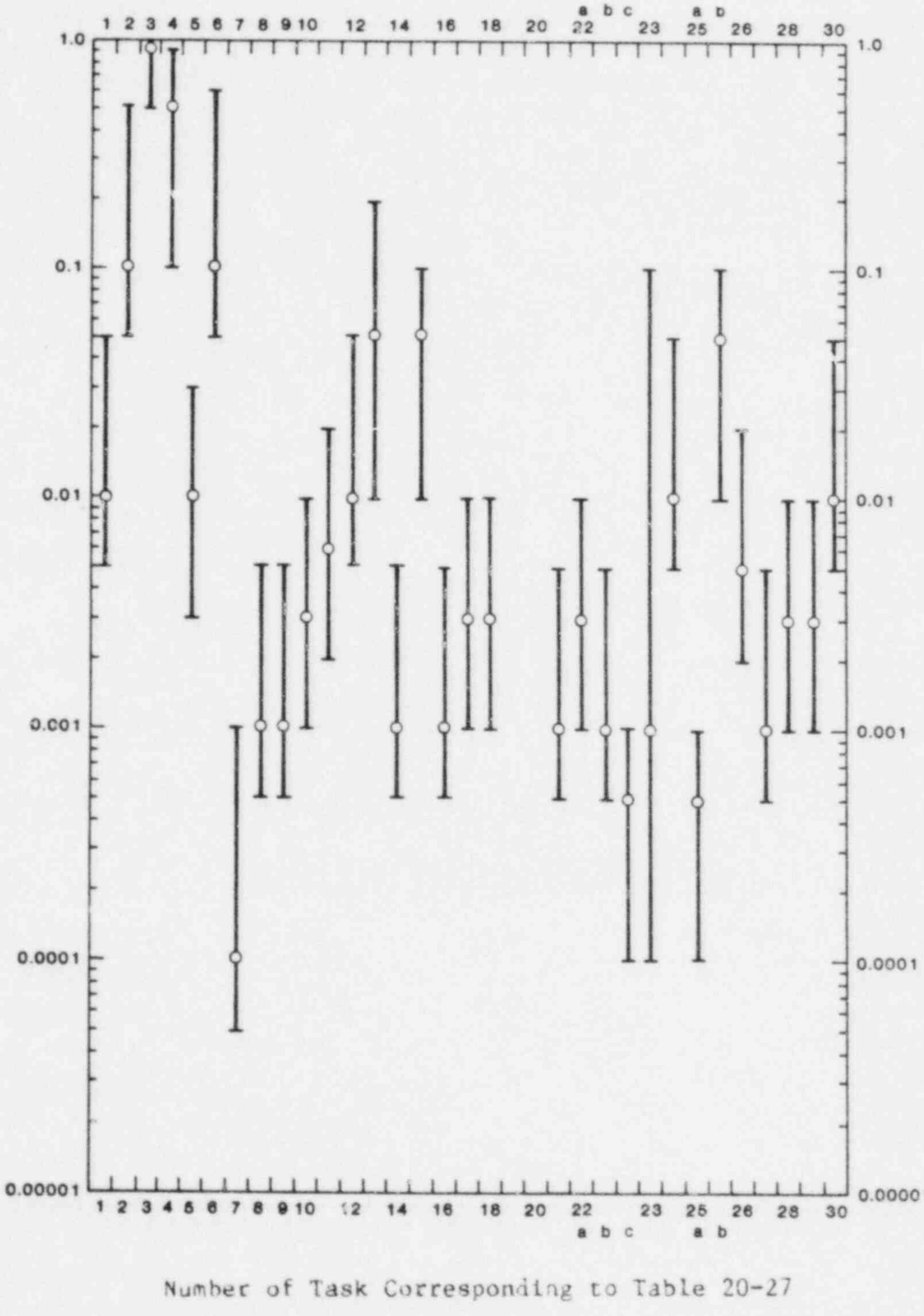
---

*The estimates in this table apply to experienced personnel. The performance of novices is discussed in Chapter 18.

**The lowest lower bound is 5×10^{-5} and the highest upper bound is .999.

Graphic Representation of HEPs

The accompanying chart and its associated table present an overview of the HEPs and uncertainty bounds associated with a number of typical tasks (Figure 20-1 and Table 20-27). Only a sample of tasks is outlined.



Number of Task Corresponding to Table 20-27

Figure 20-1. Graphic Representation of HEPs and Error Bounds for 30 Selected Tasks

Table 20-27. Legend for Figure 20-1

	Task	HEP
	Commission Errors	
1.	Walk-around inspections: recognize incorrect status, using checklist correctly	.01 (.005 to .05)
2.	Walk-around inspections: recognize incorrect status, using checklist incorrectly	.1 (.05 to .5)
3.	Walk-around inspections: recognize incorrect status, no checklist, first walk-around	.9 (.5 to .99)
4.	Use checklist correctly	.5 (.1 to .9)
5.	Follow established policies or procedures	.01 (.003 to .03)
6.	Passive inspection	.1 (.05 to .5)
7.	Respond to an annunciator (one of one)	.0001 (.00005 to .001)
8.	Read annunciated lamp	.001 (.0005 to .005)
9.	Read digital display	.001 (.0005 to .005)
10.	Read analog meter	.003 (.001 to .01)
11.	Read analog chart recorder	.006 (.002 to .02)
12.	Read graph	.01 (.005 to .05)
13.	Read printing recorder (cluttered)	.05 (.01 to .2)
14.	Record more than 3 digits	.001 (.0005 to .005)
15.	Detect a deviant meter with limit marks during initial audit	.05 (.01 to .1)
16.	Check-read specific meters with limit marks	.001 (.0005 to .005)

(Table continued on next page.)

Task	HEP
Commission Errors	
17. Check-read specific meters without limit marks	.003 (.001 to .01)
18. Check wrong indicator lamp in a group of similar lamps	.003 (.001 to .01)
19. Note incorrect status of an indicator lamp (in a group)	.99 (.98 to .998)
20. Note incorrect status of a legend lamp (in a group)	.98 (.96 to .996)
21. Remember oral instructions, one of one	.001 (.0005 to .005)
22. Select wrong panel control:	
a. Among a group of similar controls	.003 (.001 to .01)
b. If functionally grouped	.001 (.0005 to .005)
c. If part of a mimic-type panel	.0005 (.0001 to .001)
23. Set a multiposition switch	.001 (.0001 to .1)
24. Mate a connector	.01 (.005 to .05)
25. Turn control in wrong direction:	
a. If no violation of populational stereotype	.0005 (.0001 to .001)
b. If populational stereotype is violated	.05 (.01 to .1)
26. Select manual valve from a group of similar valves	.005 (.002 to .02)

(Table concluded on next page.)

Task	HEP
Omission Errors*	
27. Each item on a short list**, using checkoff	.001 (.0005 to .005)
28. Each item on a long list, using checkoff	.003 (.001 to .01)
29. Each item on a short list, not using checkoff	.003 (.001 to .01)
30. Each item on a long list, not using checkoff	.01 (.005 to .5)

---

*Omission errors include steps in any kind of procedure: valve operations, switching operations, locking of valves, etc.

**Short list = 10 items or less; long list = more than 10 items.

PART V. APPLICATION OF THE HANDBOOK AND CONCLUDING COMMENTS

PART V. APPLICATION OF THE HANDBOOK AND CONCLUDING COMMENTS

This part of the handbook consists of two chapters: Chapter 21, "Examples and Case Studies," and Chapter 22, "Concluding Comments."

Chapter 21 presents some applications of the THERP technique and of the human performance models and estimated HEPs. Since our models are scenario-oriented, the examples should show the user what kinds of decisions and judgments must be made when performing a human reliability analysis. The examples include actual studies made at specific plants. The technique of bounding analysis, used to derive upper and lower bounds of estimates of human-initiated failures across all plants, is presented with hypothetical examples.

Chapter 22 includes our comments on the state of human reliability analysis and on what is needed to improve the accuracy of this technology.

CHAPTER 21. EXAMPLES AND CASE STUDIES

This chapter presents examples of the application of the models, estimated HEPs, and uncertainty bounds in human reliability analysis. Although several such examples appear throughout the handbook, the ones in this chapter are presented in some detail so that the user can follow the various rationales we employed in actual studies. These studies will show that the use of the models and HEPs in a human reliability analysis is not as delimited as the steps in a cookbook.

Five studies are described:

1. A recalculation of a WASH-1400 human reliability analysis using HEPs from the handbook
2. An evaluation of the effectiveness of the LOCA procedures at a plant
3. A case study of the availability of auxiliary feedwater at plants where manual switchover from main to auxiliary feedwater is required
4. A retrospective "prediction" of the probability of detecting the two status lamps at TMI which indicated the unavailability of auxiliary feedwater
5. A bounding analysis based on the IEEE Myrtle Beach Conference of December 1979

For convenience, reference is made primarily to Chapter 20 for HEPs, since it summarizes most of the data in the handbook.

A WASH-1400 Analysis

This example is taken from pp III-67 to III-69 of Appendix III of WASH-1400, which describes "a sample human reliability analysis." That analysis is repeated here with the original calculations, and with new calculations based on the HEPs from this handbook. It will be seen that the final answers are very close, indicating that the analysis is relatively insensitive to the differences between assumptions made in WASH-1400 and the revised ones made here.

The analysis is concerned with the human reliability in shifting from the injection mode to the recirculation mode some 20 to 30 minutes after the occurrence of a large LOCA at a PWR. In the example, this shift-over has to be done manually. If it is not done correctly or on a timely basis, the consequences could be very serious, inasmuch as the pumps required for long-term cooling could be destroyed by attempting to pump down an empty RWST. The coolant in the RWST is used in the initial injection mode to keep the reactor covered. Before this coolant is completely depleted, it is necessary to perform the manual actions described below to pump water from the containment sump and recirculate it through the reactor vessel.

The analysis is based on paragraphs 4.8 and 4.9 of the PWR's written procedure entitled "Loss of Reactor Coolant." The two paragraphs are:

4.8 When the RWST reaches the low level setpoint (14.5%) and CLS (Consequence Limiting System) initiation has been reset (RESET PERMISSIVE $\leq$ .5 psig) complete the following actions:

  4.8.1 Open MOV-860A and B, suction to the low head SI (Safety Injection) pumps from the containment sump.

  4.8.2 Stop the containment spray pump motors and close spray pump turbine steam supply valves MS-103A, B, C, and D.

  4.8.3 Close spray pump suction and discharge valves MOV-CS-100A, 100B, 101A, B, C, and D.

4.9 When the RWST reaches the low-low level setpoint (7%) complete the following actions:

4.9.1 Close MOV-862, suction to the low head safety injection pumps from the RWST.

4.9.2 Open the charging pump suctions from the discharge of the low head pumps by opening MOV-863A and B.

Our analysis is restricted to Steps 4.8.1, 4.9.1, and 4.9.2. The MOV switches involved are MOV-1860A and B, MOV-1862, and MOV-1863A and B. (NOTE: The utility's written procedures drop the initial digit since it is understood, for example, that MOV-860A could refer to this switch for either the Number 1 reactor (i.e., MOV-1860A) or the Number 2 reactor (MOV-2860A)). These switches for the Number 1 reactor are shown in the bottom row in Figure 21-1. The two rows of switches shown are the bottom two of seven rows on the left-most panel of four segments of a large switchboard.

There are two indicator lamps above each switch: G stands for green (closed condition of MOV or pump stopped), R stands for red (open condition of MOV or pump running), and Y stands for yellow (an intermediate condition). Before the low level setpoint is reached, MOV-1862 is open (red lamp) and MOV-1860A and B and MOV-1863A and B are closed (green lamps).

Not shown in the sketch, but of importance to the analysis, is the third row from the bottom of MOV switches. This row consists of five switches identical in shape, size, and arrangement to the switches on the bottom row. The five switches are labeled from left to right as follows (continued p 21-5):

LO HEAD S.I. PP A DISC ISO VV
MOV-1864A
ISO DISC FROM COLD LEGS

LO HEAD S.I. PP A RECIRC ISO VV
MOV-1885A

LO HEAD S.I. PP A&B RECIRC ISO VV
MOV-1885C

LO HEAD S.I. PP B RECIRC ISO VV
MOV-1885B

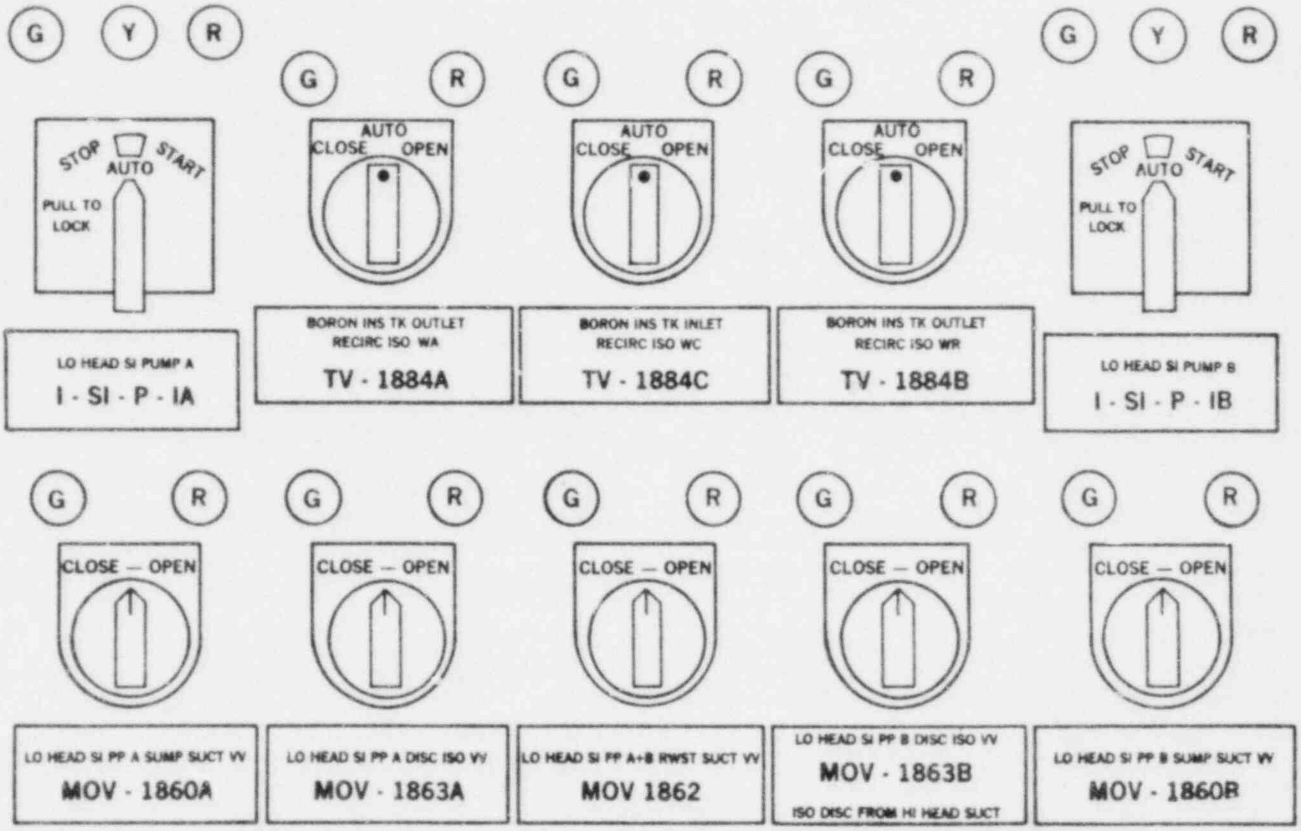


Figure 21-1. MOV Switches on Part of the ESF Panel at the PWR Used in the WASH-1400 Study (Note: The sketch is also Figure 3-4, and is based on Figure III 6-2 from WASH-1400.)

LO HEAD S.I. PP B DISC ISO VV
MOV-1864B
ISO DISC FROM COLD LEGS

(The red lamps are lit, indicating
the normally open conditions of the
valves)

It is assumed that the low level setpoint (14.5%), will be reached about 20 to 30 minutes after a LOCA (in WASH-1400, 30 minutes was assumed). When the low level setpoint is reached, all actions called for in paragraph 4.8 must be accomplished within 2 minutes, because the low-low level setpoint (7%) may be reached by then, and the operators must be ready to take the actions called for in paragraph 4.9. The approach of each setpoint is indicated by meters that display the water level in the RWST. Annunciators sound when the setpoints are reached.

Two questions were addressed in the analysis:

(1) What is the probability that no action would be taken at the low level setpoint (an error of omission)?

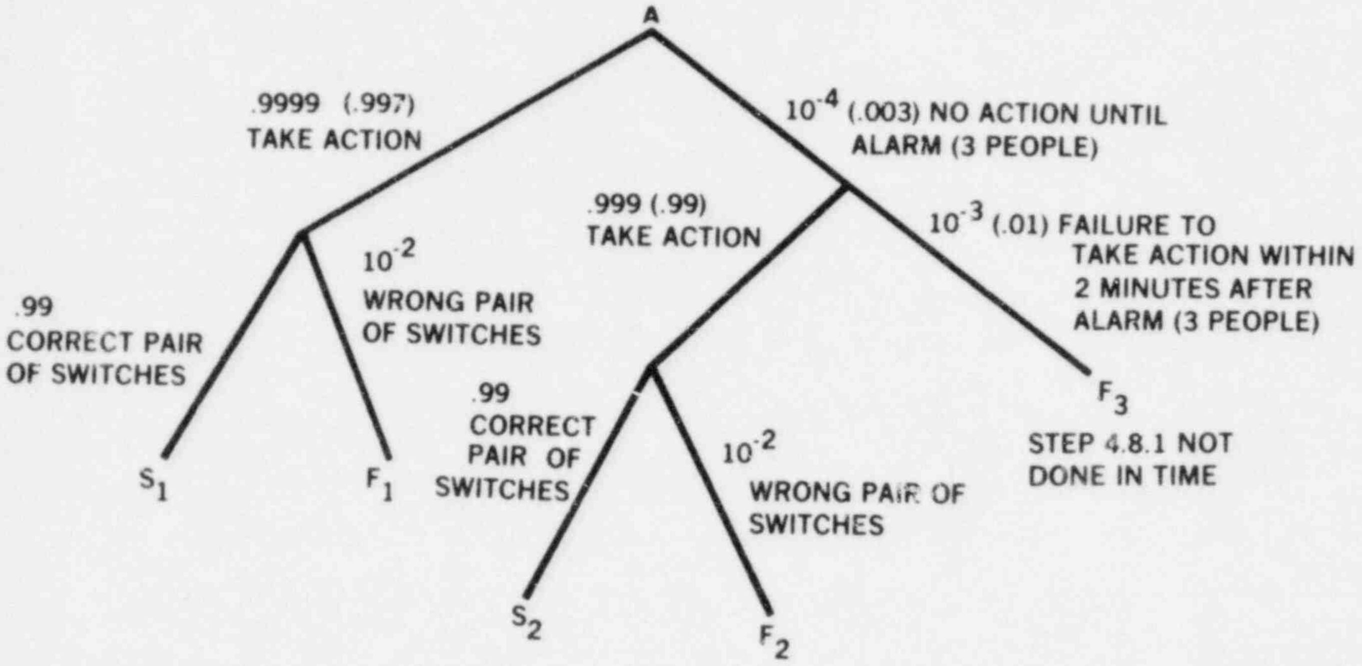
(2) What is the probability that some pair of switches other than MOV-1860A and B would be manipulated (an error of commission)?

In addressing the first question, Table 20-25 (p 20-34) indicates a basic operator HEP of 10^{-1} 30 minutes after a LOCA. It was assumed (as shown now in Table 20-24, p 20-33) that at least three people would be present in the control room by 30 minutes after a LOCA, and that action would be taken promptly unless all three failed to anticipate the low level setpoint. It was judged that the meter indication of a falling RWST level--a planning cue --should remind those present to get ready to perform Step 4.8.1. If no preparations were made before the annunciator alarmed at the low level setpoint, the chances of their completing the procedure correctly in the required 2 minutes would be greatly reduced.

In the WASH-1400 analysis we estimated that each of the three people would have a .5 probability of failing to notice the meter indications. Our rationale was that, although they know that during a large LOCA the coolant from the RWST is being used up, under the stressful condition of many alarms, danger to the plant, and potential danger to the environment, the best one could expect is about a 50-50 chance per person that he would remember to monitor the RWST level meter before the annunciator alarmed. At the PWR studied, the meter is located on a vertical panel several feet behind the desk-type panels on which the switches for Step 4.8.1 are located. This is obviously not an ideal location for a cue to remind someone to perform an action. The estimated HEP of .5 was our judgment based on this situation; no other data related to this task were available.

In the original analysis we judged the joint probability that all three people would fail to notice the meter indication to be $.5^3 = .125$, rounded to .1. We also judged the joint probability that all three people would fail to prepare for the low level setpoint as the BHEP of .1 cubed, (10^{-3}). Thus, the joint probability of failure to anticipate the low level setpoint and failure to notice the meter readings was taken as $.1 \times 10^{-3} = 10^{-4}$. This 10^{-4} is one of the values assigned to the first failure limb in the event tree in Figure 21-2. The value in parentheses is our present estimate, as discussed next.

Using our present modeling, we still assume three operators present at 30 minutes after the LOCA, but instead of the originally assumed ZD between operators, we now assume MD between the first two operators and HD between the second and third operators (Table 20-24, p 20-33). Our rationale is that the operator on duty when the LOCA occurred would be the regular operator, the second operator would be the shift supervisor, and that there is MD between these two. To be conservative, we assume that the third operator



NOTE: Where our present estimates differ from the original estimates in WASH-1400, our present estimates are shown in parentheses.

Figure 21-2. Probability Tree Diagram for Step 4.8.1 in LOCA Procedure (taken from Figure III 6-3 in WASH-1400) would be either a novice or an auxiliary operator, and we judge that there would be HD between him and either of the other two.

To calculate the first failure limb in Figure 21-2, we take the HEPs of .5 and .1 and apply the dependence equations, p 20-5, as follows:

Pr[Failure of all three operators to notice meter and failure to prepare]

$$= [.5 \times \frac{1 + 6(.5)}{7} \times \frac{1 + .5}{2}] \times [.1 \times \frac{1 + 6(.1)}{7} \times \frac{1 + .1}{2}]$$

$$= .00269 \approx .003$$

In Figure 21-2 the .003 value is shown in parentheses following the original WASH-1400 estimate of 10^{-4}.

When the low level annunciator has sounded, the operators have only 2 minutes to perform the steps in paragraph 4.8. In WASH-1400 it was reasoned that if no action had been planned until the alarm sounded, some degree of disorganization was indicated, and the BHEP of .1 was assigned to each of the three operators. A joint probability of $.1^3 = 10^{-3}$ was estimated for the failure of all three operators to take action within 2 minutes after the auditory alarm at the low level setpoint. This probability is shown on the second failure branch in Figure 21-2 leading to failure event F_3 (failing to perform step 4.8.1 in time).

Our new estimate for this branch is .01, derived by using the dependence equations as above. The values for F_3 using the WASH-1400 estimates and our new estimates can now be compared. The original estimate was $10^{-4} \times 10^{-3} = 10^{-7}$. It was stated in WASH-1400 that any estimated HEP of less than 10^{-5} should be viewed with skepticism, but it could be concluded that the probability of failure to perform Step 4.8.1 was relatively small. This potential failure sequence was therefore dropped from further consideration. With the new estimated HEPs, $F_3 = 3 \times 10^{-3} \times 10^{-2} = 3 \times 10^{-5}$, still a small enough number as to have no major effect on the overall analysis. However, the new figures do provide a more conservative estimate of the human reliability for Path A → F_3.

The next step in the original analysis was to assume that at least one of the three operators did prepare to manipulate switches MOV-1860A and B. This leads to the second question: What is the probability that some pair of switches other than MOV-1860A and B would be manipulated? This is represented in two places in the event tree--in the terminal limbs leading to F_1 and F_2. The CHEP for this task was estimated as 10^{-2}. Our rationale was that it would be highly probable that responsibility for operating the valves would be assigned to one person, that is, no human redundancy would be available to recover an error by the person performing this task. This judgment was based on observation of operators at work. Misselection of switches is the type of error that few operators regard as a credible error. Therefore, it was deemed unlikely that anyone would check the operator who actually manipulated the MOVs. The basic error probability of 10^{-1} was assessed to be too large for this act, so 10^{-2} was selected as the nearest order-of-magnitude estimate. We have no reason to change the original reasoning and estimate.

Now it is possible to assign HEPs to all of the limbs, since the sum of the probabilities at each branching must equal 1.0. Two paths lead to misselection of the switches: Path A → F_1 and A → F_2. Using the original estimates, the probability of going down the second path is:

$$\text{Path A} \rightarrow F_2 = 10^{-4} \times .999 \times 10^{-2} \cong 10^{-6}$$

In WASH-1400 this small probability was eliminated from further consideration.

Using the new estimates the probability of this path is $.003 \times .99 \times .01 \cong 3 \times 10^{-5}$. In most circumstances this value would also be excluded from further consideration.

Using the original estimates, the probability of Path A → F_1 is .9999 x 10^{-2}, which rounds to 10^{-2}. The comparable probability using the new estimates is .997 x 10^{-2} which also rounds to 10^{-2}.

The rest of this section continues the WASH-1400 analysis to illustrate further applications of human reliability analysis. As can be seen, the HEP estimates were based on judgment, as no "hard" data were available.

Given that the operator selects a wrong pair of switches at the low level setpoint the question arises as to which incorrect pairs might be selected. It was judged that the most probable candidates are MOV-1863A and B: they are in the same row, are next to the desired switches, and have similar MOV numbers and labels. The probability of selecting a pair of switches from the second row from the bottom is lower because of the dissimilarity of switch nomenclature and the different appearance of the switches themselves (they have an AUTO position). The switches in the third row from the bottom have labels similar to those on the desired switches, but the outboard switches (the most likely candidates for misselection) are normally open. Their red indicator lamps would furnish a cue that they are not the correct switches. In addition, this third row is somewhat remote from the desired switches.

Given the initial error of selecting some pair of switches other than MOV-1860A and B, it was estimated that there was a probability of .75 that the operator would select MOV-1863A and B, and a probability of .25 that some other pair of switches would be selected. (The .75 and .25 split was assessed on the basis of the layout of the switches, and represents the kind of judgment that is independent of the HEPs in the handbook.) The error of misselection of MOV-1863A and B has a recovery factor at the 7% (low-low level) setpoint. In Step 4.9.2 the operator is supposed to close MOV-1863A and B. If the error of misselection had been committed, the operator would find these MOVs already closed. This should cue him that something is wrong. An HEP of .1 is assumed for failure to note the error. Hence the total estimated failure probability for Step 4.8.1, including failure of the recovery factor, is $10^{-2} \times .75 \times 10^{-1} = .00075$, which is rounded to 10^{-3}. (The HEP of .1 is the HEP assumed for most operator actions 30 minutes after a large LOCA, as shown in Table 20-25, p 20-34.)

A similar analysis was performed for Steps 4.9.1 and 4.9.2. The detailed analytical approach described above involved some degree of subjectivity. This subjectivity was not particularly crucial for the study because the important factor that affects overall results is the order of magnitude of the human error probability, not its exact value. The error bounds assigned in WASH-1400 to the final estimate allowed for uncertainties and errors in the analysis. As a tool in itself, the detailed analytical approach is valuable for the following reasons:

a. The exercise of outlining all plausible modes of operator action decreases the probability of overlooking some important failure path.

b. Due to the lack of error probability data for nuclear power plant tasks, it is necessary to break down operator actions to a level at which existing data can be used.

c. The detailed approach makes it easier for analysts making independent estimates to check on the source of any disagreement and to resolve it.

Case Study of a LOCA Procedure

This example is based on a study performed for the NRC as part of their continuing evaluation of emergency procedures. The study was primarily a qualitative analysis, and we have included it to show how the handbook can be used for qualitative as well as quantitative assessments.

In this plant (a PWR), human participation would be required early in the post-LOCA sequence of recovery actions. The most critical human action would be that of switching from the Safety Injection (SI) of emergency coolant to the recirculation of coolant from the sump to the reactor. This action is similar to the action described in the previous example, except that the switching task would have to be done by about 8 minutes after the LOCA rather than the 30 minutes assumed in the first example. The worst-case analysis is based on the assumption of the same type of guillotine break in a large coolant pipe as assumed in WASH-1400, with further assumptions that the plant is at full power and that the water level in the RWST is at the lowest level allowed by the technical specifications; i.e., a design-basis LOCA.

Reference to Figure 17-2 (p 17-20) shows that at about 8 minutes into a large LOCA an operator would be performing correctly only one time in four. Rather than rely on this speculative curve, we conducted an informal experiment in the plant in which the shift supervisor walked through all the actions called for by the written operating instructions. The shift supervisor was considered the best operator at the plant, yet starting from a "go" signal he was barely able to complete the procedures in 8 minutes. This ideal response situation did not allow for any time lost due to indecision about the existence of a LOCA, nor was any allowance made for secondary failures of instrumentation in conjunction with a LOCA. For example, if the containment sump level indicating lamps failed, the first warning to the operator of the need for changeover to the recirculation mode would occur at 471 seconds (nearly 8 minutes) after the LOCA, the second would occur at 500 seconds (8-1/3 minutes), and a third at 503 seconds. If the operator failed to initiate switchover at the third indication, all of the available coolant would be gone in another 92 seconds and the feedwater pumps would be damaged.

During these 92 seconds he must perform eight switching actions at three panels several feet apart.

We concluded that we could not be confident that the switchover would be accomplished in time in the event of such a design-basis LOCA. We identified the basic human factors problems as those of time stress, poorly written procedures, and poor human engineering of the control room. Specific suggestions were made to reduce the effects of these three problems on human reliability.

In the first area, time stress, some reduction had already been achieved at the plant by eliminating some of the decision-making required of the operator in determining whether a LOCA had occurred. Their training program directed an operator to assume a LOCA when certain annunciated symptoms appeared (pressurizer low level, pressurizer low pressure, containment high pressure, etc). We recommended that talk-throughs like the one we conducted be held frequently, and that the onsite NRC inspector periodically check the readiness of an operator to respond.

In the second area, written procedures, we noted that a number of operator actions were required primarily for economic considerations and not for the switchover. We recommended that all such steps be postponed until the switchover had been completed. This would considerably reduce the number of actions an operator would have to take in the critical first 8 minutes. We further suggested changing the sequence of written steps in the procedure to reduce travel between different panels--in link analysis terms, to reduce the number of required links (Figure 3-3, p 3-29). Finally, we suggested rewriting the procedures in the columnar format shown in Figure 3-12 (p 3-51).

In the third area, human engineering, we suggested relabeling displays and controls in a consistent fashion, using the type of location aid shown in Figure 3-6 (p 3-33). With these three changes, and the NRC inspector's frequent verification that operators selected at random could indeed carry out the procedure, we estimated that substantial gains in human reliability would result. We did not attempt to quantify the projected human reliability at the time, as we felt that such an evaluation could be carried out more meaningfully after the changes had been implemented.

Case Study of Manual Switchover to AFWS

As part of an NRC post-TMI study of the availability of the AFWS, we visited a PWR where the switchover from main feedwater to aux feedwater is done manually. Plans were under way at this plant to incorporate automatic switchover, and the question was whether manual switchover was a safe interim procedure, not only at this plant but also at other plants. At some plants, the switchover must be made within 5 minutes or the steam generator might run dry. In other plants as much as 15 minutes is available, and at still others 30 minutes is available. We were asked to prepare human reliability estimates for manual switchover under conditions which allowed 5, 15, 30, and 60 minutes to accomplish the switchover. (The interest in the 60-minute period was not related to this particular problem.)

At the plant we visited, a second operator whose only function was to maintain sufficient water inventory in the event of a transient was assigned to the control room. This operator was designated a dedicated operator (DO), as distinct from the regular reactor operator (RO), who monitored the rest of the control room. The DO at this plant was required to stay within a small designated area of the control room. Because this job is confining and not very demanding, the plant assigned operators who had not yet achieved status as fully qualified ROs, but who wanted to achieve this status. It was judged that their motivation would be sufficient for this confining job. All DOs recognized that this was an interim assignment (until the automatic provisions could be incorporated) and that management took notice of their willingness and ability to serve.

The plant had also adopted a procedure to eliminate the need to decide to initiate the AFWS. This procedure called for the DO to switch from main to aux feedwater *whenever* a reactor trip occurred if the plant was operating at more that 15% power. This plant policy had important bearing on our HEP estimates.

The switchovers from main to aux feedwater were performed frequently at the plant, in both real and simulated situations, so that the manual actions involved in switchover were very well learned by the DOs. Therefore, we could assume that the only significant source of human error would be the failure to begin the switchover procedures.

On the basis of interviews with and observations of operators at the plant, we estimated the probabilities of errors of oversight for plants with and without a DO. Our original estimates are presented in Table 21-1 along with our revised estimates. The revised estimates are based on changes to the dependence model made subsequent to the original analysis. The discussion below presents our rationale for the original estimates and the subsequent changes.

We will first consider the situation without a DO, in which the RO would have to initiate AFWS in addition to his other duties. The HEP of .05 for the first 5 minutes is drawn from the model for annunciated displays (Table 20-4, p 20-10). In the event of the need for AFWS, 40 or more annunciators may be sounding, and the RO has to integrate all the information and make appropriate decisions. Note that Table 20-4 indicates an HEP of .25 for 40 or more annunciators. Our rationale for using the HEP of .05 instead of .25 is based on the following: in plants requiring manual changeover to AFWS,

Table 21-1. Estimated Probabilities of Failure to Initiate AFWS*

At end of X min	Situation without Dedicated Operator				
	Regular Operator	Shift Supervisor Orig. HEPs	Shift Supervisor New HEPs	Total Failure Probability Orig. HEPs	Total Failure Probability New HEPs
5	.05	-------	------	.05	.05
15	.01	.5 (MD)	.5 (HD)	.005	.005
30	.005	.25 (LD)	.15 (MD)	.001	.001
60	No change	No change	No change	No change	No change

	Situation with Dedicated Operator				
	Dedicated Operator	Shift Supervisor Orig. HEPs	Shift Supervisor New HEPs	Total Failure Probability Orig. HEPs	Total Failure Probability New HEPs
5	.002	-------	------	.002	.002
15	.001	.5 (MD)	.5 (HD)	.0005	.0005
30	.0005	.25 (LD)	.15 (MD)	.0001	.0001
60	No change	No change	No change	No change	No change

*Lower and upper uncertainty bounds of a factor of 10 are assigned to each estimate in the "total" columns.

All HEPs are rounded.

plant policy usually requires the changeover to be initiated any time a turbine/reactor trip occurs. Although each of the annunciators conveys a unique message, the messages are not independent, and we estimate that at least 10% of them would convey indications that a trip has occurred. The HEP of .05 is the estimated probability of failure to initiate action in response to one randomly selected annunciator of ten. If plant policy did not mandate the switchover to AFWS any time a turbine/reactor trip occurred, we would use the .25 estimated HEP associated with 40 or more annunciators.

We judged that if the time constraints were loosened from 5 to 15 minutes, the probability of the RO's failure would be reduced by a factor of 5; hence, the .01 estimate. With about 30 minutes, we allowed another factor of 2 reduction in the HEP, to .005. (These estimated HEPs represent judgments not included in the tabled values in the handbook.)

The RO's failure to initiate AFWS on time could be compensated for by the shift supervisor as a backup. We assumed that for the first 5 minutes the shift supervisor would not be available as a backup operator in the control room (Table 20-24, p 20-33). For the period between 5 and 15 minutes, we estimated that he would be available but still "coming up to speed." In our original analysis we estimated that the conditional probability of the shift supervisor's failure to compensate for the RO's failures was .5. In the old dependence model this .5 estimate was equivalent to a moderate level of dependence between the shift supervisor and the RO. We still believe the .5 CHEP. In the new dependence model, this is equal to the assessment of a high level of dependence. Then, as now, we estimated the conditional probability of failure directly, without invoking the dependence model until after the fact. Generally, we prefer to estimate conditional probabilities directly rather than rely on the dependence model. (Of course, if there were better data available, we would use them.)

For the period up to 30 minutes, we reduced the shift supervisor's HEP by a factor of 2 to .25, which corresponded to an assignment of LD in the old dependence model. We see no reason to modify this estimated reduction, but in the new estimates we have shown the nearest HEP with the new dependence model, for illustrative purposes. This is .15, which corresponds to MD. The reduction of the CHEP for the shift supervisor from .25 to .15 makes no difference in the overall assessment (the "total" columns).

Note that we did not estimate any further improvement at 60 minutes. It was our judgment that if the AFWS had not been turned on by the end of 30 minutes, the operators would be heavily occupied and performance on the AFWS task would not improve until things were under control.

For the situation with the DO, we started with the oral instructions model (Table 20-18, p 20-28). We considered the standing instruction for the DO to initiate AFWS when there is a reactor trip to constitute one oral instruction, and the HEP assigned to this is .001, which we assigned to the 15-minute time period. We doubled it for the first 5 minutes because we judged that for part of those 5 minutes the RO would still be subject to the incredulity response and the general reaction to many alarms' sounding in a short time. For the 30-minute period, we reduced the BHEP by a factor of 2, to .0005. The backup estimates for the shift supervisor are the same as for the situation with the RO, and the values in the "total" columns are calculated as before.

Our rationale for the low estimated HEP for the DO was based on the practices at the plant we visited. First, his duties are limited. He is responsible for only a very small portion of the control boards. Second, he performs periodic talk-throughs of the procedures involved in shutting down the main feedwater and initiating the auxiliary feedwater. Third, he has standing instructions to initiate the AFWS in the event of a turbine/reactor trip. Thus, he has very little interpretation to do and his response should be almost automatic. On the basis of this qualitative analysis, we decided that the oral instructions model was appropriate.

Note that our estimates are pegged at the ends of certain time intervals. We did not attempt to plot a curve of HEPs vs time, as this would have been pseudo-precision in view of the subjective nature of the estimates.

The Two Status Lamps at TMI

Part of the problem in the TMI accident was that two blocking valves had not been returned to their normally open status after maintenance. These two valves are MOVs controlled by two switches designated as EF-V12A and B. Each has two status lamps (red for open and green for closed). The switches were positioned on the control panel with EF-V12B immediately above EF-V12A. A yellow caution tag attached to an instrument above EF-V12B covered both status lamps of that switch. The fact that one of the pair of lamps was covered would have made no difference, as the two lamps comprise one completely dependent set.

Some people have expressed surprise that control room personnel did not detect the wrong state of the status lamps during the five shifts up to and including the shift in which the accident occurred. (Stello, 1979, p IA-18, surmises that the valves were not restored after a surveillance test at approximately 10 a.m. on March 26. The demand for aux feedwater occurred at 4 a.m. on March 28.) We estimate a .99 probability of failure to detect a status lamp in the incorrect state during the initial audit (Table 20-10, p 20-15), assuming that the initial audit includes the requirement to scan the given status lamps. If the original error of forgetting to restore the valves occurred at 10 a.m. (i.e., after the initial audit on March 26 for the day shift), there would have been four initial audits in which this anomaly could have been detected. The estimated probability of such a detection is only $1 - .99^4 = .04$.

In actual practice, the status lamps in question are not checked on any routine basis. According to operating personnel we have interviewed, it is assumed that the lamps are in the correct position unless something causes operators to question this assumption. There are hundreds of such lamps that display static status.

We can reasonably conclude that it would have been highly unlikely for the incorrect status of these lamps to have been noticed during the routine operations in the control room. (The causes for the original error of failing to restore the blocking valves after maintenance are another matter.)

Bounding Analyses

Often estimates of certain potential human-initiated failures are required even when there is very little specific information to work with, and the estimates are to be applied to NPPs in general. (This happens despite the wide variability in PSFs across plants.) Although only gross estimates can be developed in such cases, it is possible to establish a range of probabilities with fairly accurate limits to the range. The method used is called a bounding analysis. In performing a bounding analysis, one addresses some set of potential failures affected by human errors and determines an upper and lower failure probability bound for each failure. Typically, the upper bound is the 95th percentile bound and the lower bound is the 5th percentile bound, so that the probabilities assigned to these two bounds include 90% of the failure probabilities for each event. The output of a bounding analysis is generally a statement that, for a given failure event, the analyst is confident that the probability of that event is no lower than some value and no higher than some other value.

Each potential failure event is generally made up of several human tasks (or actions). For example, one failure event might be the probability that, after a surveillance test, shift personnel fail to restore a set of valves related to some ESF, leaving it unavailable. Chapter 13 indicates that there are many possible scenarios for the human errors and recovery factors related to valve restoration, and that each scenario has different combinations of PSFs that affect the estimates of human errors.

There are different methods for estimating "no lower than" and "no higher than." The general approach is to prepare an event tree with all the steps and PSFs relevant to the failure event, listing the nominal HEPs and the ranges of the HEPs for each item. For the best-case analysis one uses the lower bounds of the HEPs and the upper bounds of the HSPs associated with the success paths through the tree. The terminal values of the success paths are summed and subtracted from 1.0 to obtain the lower bound of the overall failure probability. Obviously, this lower bound for the system no longer corresponds to the 5th percentile bound. For the typical analysis consisting of several events, it can be considered as a system bound much lower than the 5th percentile bound. For the worst-case analysis one uses the upper bounds of the HEPs and the lower bounds of the HSPs associated with the failure paths through the tree. The terminal values of the failure paths are summed to obtain the upper bound of the overall failure probability. As is the case with the lower bound, this upper bound does not correspond to the 95th percentile bound. For the typical analysis consisting of several events, it can be considered as a system bound much higher than the 95th percentile bound. If no other variables are to be evaluated, the same results will be obtained whether all calculations are performed on the success paths alone or on the failure paths alone. However, by calculating the success and failure paths separately, different assumptions can be evaluated more conveniently.

To test for the effects of individual tasks on overall bounds, assumptions may be introduced regarding their HEPs, as in a sensitivity analysis. For example, if one of the PSFs in an analysis is "the proper use of procedures," and the effect of this PSF is to be estimated, one can perform the analysis with the optimistic assumption that all personnel will use written procedures correctly, and assign an HEP of zero to this item. The analysis can be recalculated with the pessimistic assumption that none of the personnel will use the written procedures correctly, assigning an HEP of 1.0 to the item. Such variations will be described later.

Specimen Tasks for Bounding Analyses

At the IEEE Standards Workshop on Human Factors & Nuclear Safety held at Myrtle Beach, SC, on December 2-7, 1979 (Schmall, 1980), a number of tasks were submitted to the conferees as exercises in estimating upper and lower bounds for HEPs. The information was incomplete and, in some cases, ambiguous. Such inadequacies in task description are common in preliminary reliability analyses. We have selected three specimen tasks from the IEEE exercise, and will develop a bounding analysis for each. The three tasks are:

1.3 - Sampling or Valve Test Activity

1.4 - Maintenance, Pump, or Component Activity Involving Two Valves

1.5 - Valve Line-Up Activity

In the following descriptions of the tasks we use the wording of the original material, omitting only some irrelevant or misleading sentences. Also, we substituted the term "error probability" for "error rate" in keeping with the terminology used in this handbook. With these exceptions, the following material between the two sets of asterisks is taken verbatim from the IEEE Workshop instructions:

* * * * * * * * * * * * * *

1.0 HUMAN ERRORS ASSOCIATED WITH MANUAL VALVES

1.1 GENERAL INTRODUCTION OR GROUND RULES

During all phases of nuclear power plant operation, the human interacts with many different types and sizes of manual valves. These valves have the following characteristics in common:

1. They require the person who is performing the operation to be directly at the site of the component.
2. They range is size from small (1/4" to 4") which usually are direct acting or require no mechanical assistance for valve motion, to large (up to 20") which usually have gearing or leverage devices to assist the human or magnify the force he applies for valve motion.
3. They are located in piping throughout the plant (not in the control room or other specific location).
4. They are not instrumented with position indicators and are not capable of signaling position or motion except at their specific location, although secondary indications such as process pressure, temperature or flow may be interpreted as position for certain specific valves.

It is assumed that in all cases, the activity is carried on by equipment operators or maintenance workers, who have less training and systems knowledge than licensed reactor operators, but presumably more experience with operation of individual manual valves.

In all cases, it is assumed that the manual valve involved was in the correct position prior to the activity indicated, and was left in the wrong position as a consequence of this activity. Also we assume that "correct" position is the same for both normal and emergency operation, since a plant should not be designed to require movement of a manual valve during an emergency.

1.2 FACTORS CONSIDERED

With the ground rules outlined above, the major factors which are considered to potentially affect human error probabilities for manual valves are as follows:

1.2.1 Types of Activity

Type of activity, of which three significantly different types have been identified. These are sampling or valve testing; maintenance, pump, or other component testing; and operational valve lineup. The differences are outlined in paragraphs 1.3, 1.4, and 1.5 below.

1.2.2 Valve Location

Location of the valve, of which two different categories are listed. One location is "near" the equipment or component affected when more than one valve is involved in the activity. This would include valves directly on the discharge or inlet to a pump, within view when standing at the pump and easily accessible from the pump (i.e., no climbing or going into another space or enclosure is involved). Another condition required for this "near" category is the absence of environmental discomfort or hazard (i.e., no special clothing or precautions involved due to radioactivity or climatic hazard). Typically a "near" valve is located within 10 feet of its associated component (pump, heat exchanger, etc) in the same space or compartment, both located in an area with a comfortable shirt sleeve atmosphere.

The other category, "remote", is either far from the associated component (over 20 feet), located in another space (hidden from view from the component affected), and/or requires special clothing or precautions because of its location.

1.2.3 Valve Size

Valve size is considered a factor, and for simplicity has been divided into two distinct categories, small and large. Small sizes are those which can be moved directly by one man and lend themselves to grouping or manifolding; assume 2" diameter. Larger sizes usually require mechanical leverage or gearing to operate, and location is usually determined by piping layout, not accessibility; assume 18" diameter.

1.2.4 Type of Procedure

Procedure used which may be divided into written procedure or oral directive.

The written procedure may be a test, maintenance or operating procedure written order with as many as 100 steps. For our purposes, we will assume 50 steps. The steps which affect a particular valve are usually limited to one or two steps dispersed throughout the procedure. We assume here that the operator is required to check off the steps he has performed on the written procedure.

The oral directive is usually given to the equipment operator by another equipment operator or reactor operator in the course of normal plant operation. We assume here that the operator giving the directive is following a procedure known to him, but the operator accomplishing the activity does not have a written procedure to guide him. We assume in this case that there is no initialing or checking off of any written list.

1.3 SAMPLING OR VALVE TEST ACTIVITY

This activity, for our purposes, is defined to be tests of 12 valves. The test is of short duration (say, about 30 minutes), and involves a single operator from start to finish. The person involved prepares himself for the location beforehand (properly dressed and equipped), so that location factors do not affect the outcome of the activity. Typically, as for the quarterly valve tests prescribed by the ASME code, the equipment operator is sent out to various locations on a definite schedule, with a list of valves and a test procedure for each valve or group of valves. We will assume here that the test requires changing position to prove operability and returning it to the correct operating position. We assume that 12 valve tests are performed in sequence immediately after one another. The error considered is forgetting to return one or more of the valves to the correct position. The operator must record the action taken on a data sheet.

1.4 MAINTENANCE, PUMP, OR COMPONENT TEST ACTIVITY INVOLVING TWO VALVES

This activity is typified by maintenance or test of in-line components such as pumps, heat exchangers, filters, etc, and usually involves two or more valves. For our purpose we assume two valves are involved. One or more equipment operators are given a written job order or procedure and a specific schedule to accomplish the work. This usually involves valving a spare component on line (in the flow path) and valving the affected component off line (out of the normal flow path). The valves may be located within sight of the affected components, in a comfortable shirt sleeve atmosphere (near location), or they may be in another space and/or require special clothing or equipment (i.e., a ladder) in order to operate them (remote location). After completing the maintenance or test on the affected component (e.g., pump), the valves must be returned to their original "correct" operating position. We assume that a written checkoff is required of the results of the actions taken. The errors we considered are again limited to acts of omission, i.e., forgetting to return either of the two valves involved to their "correct" operating position.

In addition to location, valve size is another factor considered potentially affecting error rates.

Variations of this activity include removal and replacement of components in a flow path, or at the end of a flow path (tanks), normal maintenance such as recharging catalyst, replacing filters, etc.

1.5 VALVE LINE-UP ACTIVITY

In the normal course of operations, it is frequently required or desirable to remove from service and return to service active components such as pumps or heat exchangers. These activities usually involve re-positioning of several valves and may be scheduled such as the normal rotation of active pumps to equalize wear and usage, or they may be unscheduled such as the substitution of a leaking heat exchanger or filter. In the latter case the activity may be accomplished without a written order or procedure, usually under oral directive from the reactor operator. Again, the error of interest here is the one of omission or failure to position all the valves involved to their "correct" operating position. To be specific, we assume two manual valves are involved and the error is leaving both manual valves in the incorrect position. The error associated with the manual valves that we are examining here is usually not discovered until some subsequent action in the control room is performed (i.e., opening a valve in series with the manual valve). The location and size factors are still variables as in the previous groups, in addition to the procedure used, written or oral.

(End of IEEE Workshop Instructions)

* * * * * * * * * * * * * * * *

Bounding Analysis of Sampling of Valve Test Activity (Example 1.3)

The procedure followed in our analysis is as follows:

(1) Identify the failure event(s) of interest.

(2) List the major human actions related to the failure event.

(3) List the relevant PSFs, ignoring those likely to have only a minor effect on the estimated HEPs. (In a bounding analysis, small effects can be ignored.)

(4) Select the type of best- and worst-case analyses to use. (In these examples we will use more than one type.)

To illustrate how the bounding analysis was originally performed, we will go through the steps in our original rough draft, without ordering them in any logical fashion.

The failure event of interest is the failure to return one or more valves to the correct position after testing. Two major errors are possible: forgetting to test any one valve, and failing to return any one valve to its fully restored position. In this analysis, if a valve is not tested, restoration errors are not possible. Therefore, we can ignore the error of forgetting to test a valve.

After determining the failure event to be analyzed, we listed the relevant PSFs as we thought of them:

1. Written procedure required - assume 50 steps: whether or not it is used
2. Checkoff required for each step in the procedure: whether or not it is used
3. A data sheet used to record action (but details not specified)
4. Whether or not valves are in same perceptual area
5. Size of valves - small (2" diameter) or large (18" diameter)
6. May or may not wear special clothes
7. Single operator (not a maintenance person)
8. Initiating cue for test: a schedule

9. Whether the valves are all up, all down, or mixed when in correct position
10. Whether any valve to be tested is among other valves and, if so, whether any of the other valves are initially in the same position as the desired valve
11. Whether a valve sticks as it is returned to its normal position
12. Whether valves are locked, and type of administrative control
13. Whether each valve is changed and immediately restored, or all valves are first changed and then all restored. Assume it is plant policy to use the logical (first) procedure, and assume no tagging because of the nature of the test.
14. Quality of labeling on valves
15. Whether valve labels are the names used in the written list of valves
16. Whether any of the 12 valves is omitted from the list
17. No radiation environment

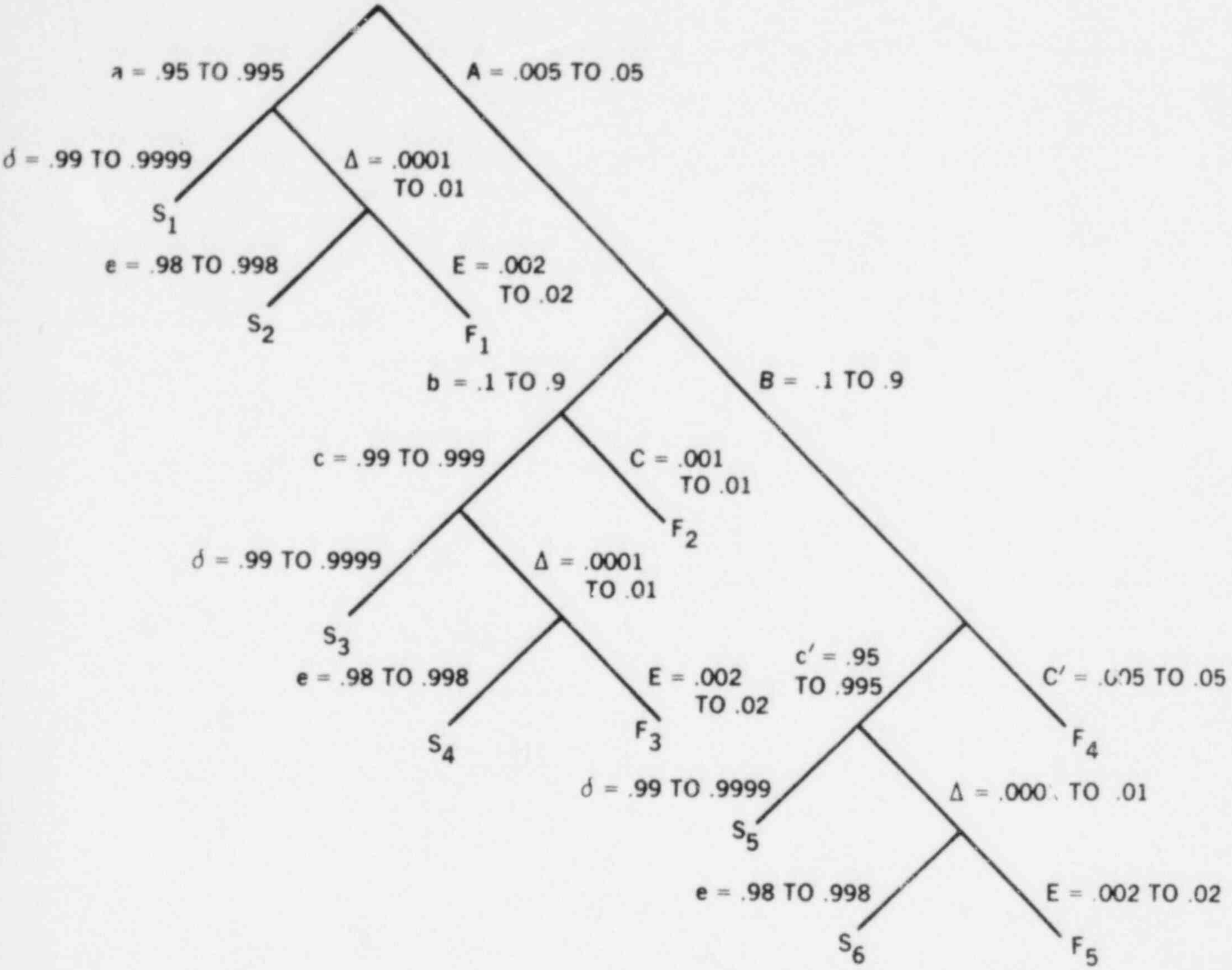
We then reviewed the list to see if there were any overriding factors, and it was clear that PSF 13 was such a factor. If the logical procedure of changing and then restoring each valve in turn was followed, we concluded that the only credible error would be failure to restore a valve <u>fully</u>. If the operator forgot a valve entirely, it would not matter for this analysis, since he could not make a restoration error on that valve. If he changed a valve, he would always restore it, since there would be no break in his sequence of rotating the valve to one position and then immediately returning it to its original position.

If the illogical procedure in PSF 13 were followed, errors of omission would be possible. For our worst-case analysis, we will assume that the operator was using the written procedures, but only as a list of valves to be tested, and for some reason was not abiding by that part of the procedures that states that each valve shall be restored before proceeding to the next. That is, the operator would change all the valves listed in the procedures, and then would have to restore those he had changed. There is no entry in the handbook for following an illogical procedure, but the entry in Table 20-22 (p 20-31) is close: Failure to carry out a plant policy when there is no check on a person. The estimated probability of this error is .01 (.005 to .05).

From Table 20-20, p 20-29, the HEP for errors of omission when using a long list with checkoff is .003 (.001 to .01), and when using a long list without checkoff, .01 (.005 to .05), the probability that the checkoff provision will be used improperly is .5 (.1 to .9). If checkoff is used improperly, the HEP and uncertainty bounds for a long list with checkoff are used.

If a valve sticks for some reason as it is restored, the HEP for failure to restore the valve completely is .005 (.002 to .02) (from Table 20-14, p 20- 21 for rising stem valves without position indicators). On the basis of component reliability estimates, the probability of a valve sticking is .001 (.0001 to .01). Assume zero probability of more than one valve sticking.

With the above HEPs, we are able to construct the event tree and determine upper and lower error bounds under a variety of assumptions. The several steps and PSFs will be designated alphabetically. To calculate the upper bound for a worst-case analysis, we calculate the sum of all the failure paths in Figure 21-3, using the most pessimistic probability bound for each limb; i.e., the upper bound of the HEP and the lower bound of the HSP. (Note: For purposes of this chapter, we are using all of the limbs in each failure path. Ordinarily, when success probabilities are greater than .95, the success limbs may be ignored; i.e., set to 1.0, in calculating the values of the failure paths.)



EVENTS	HEPs (BOUNDS)
A = FAILURE TO FOLLOW LOGICAL PROCEDURE	.01 (.005 TO .05)
B = FAILURE TO USE CHECK-OFF PROVISION PROPERLY	.5 (.1 TO .9)
C = ERROR OF OMISSION USING CHECK-OFF PROPERLY	.003 (.001 TO .01)
C′ = ERROR OF OMISSION USING CHECK-OFF IMPROPERLY	.01 (.005 TO .05)
Δ = VALVE STICKS DURING RESTORATION	.001 (.0001 TO .01)*
E = FAILURE TO RESTORE A STICKING VALVE COMPLETELY	.005 (.002 TO .02)

*TO SIMPLIFY THE CALCULATIONS, WE WILL ASSUME THAT NO MORE THAN ONE VALVE IN A GROUP STICKS WHEN BEING RESTORED.

Figure 21-3. Event Tree for Example 1.3

Note that the limb, a, leads directly to event " Δ," as we have made the assumption that errors of omission could be disregarded if the logical procedures were followed.

The five failure paths* are:

F_1. $a-\Delta-E = .95 \times .01 \times .02 =$ .00019

F_2. $A-b-C = .05 \times .1 \times (1 - .99^{12}) =$ .000568

F_3. $A-b-c-\Delta-E = .05 \times .1 \times .99^{12} \times .01 \times .02 =$ .0000009

F_4. $A-B-C' = .05 \times .9 \times (1 - .95^{12}) =$ .0207

F_5. $A-B-c'-\Delta-E = .05 \times .9 \times .95^{12} \times .01 \times .02 =$ .0000049

.0214638

Thus, the upper bound for failure ≈.02.

For a best-case analysis, we calculate the sum of all the success paths, using the most optimistic bound for each limb; i.e., the upper bound of the HSP and the lower bound of the HEP. The success paths are:

S_1. $a-\delta = .995 \times .9999 =$ .9949005

S_2. $a-\Delta-e = .995 \times .0001 \times .998 =$ .0000993

S_3. $A-b-c-\delta = .005 \times .9 \times .999^{12} \times .9999 =$ .0044459

S_4. $A-b-c-\Delta-e = .005 \times .9 \times .999^{12} \times .0001 \times .998 =$ .0000004

S_5. $A-B-c'-\delta = .005 \times .1 \times .995^{12} \times .9999 =$ .0004708

S_6. $A-B-c'-\Delta-e = .005 \times .1 \times .995^{12} \times .0001 \times .998 =$ .000000047

.9999169

Subtracting .9999169 from 1.0 yields a lower failure bound of .0000831 ≈ 10^{-4}.

---

*To calculate C or C' (errors of omission), the number of items must be considered. In this case there are 12 valves, and the task is failed if one or more are omitted. The probability of success is raised to the 12th power and subtracted from 1.0 to obtain the probability of failure.

The above analyses can be modified by making any assumptions that are of interest. For example, one might want to make the ultra conservative assumption that at least one valve will always stick, in which case the probability of Δ becomes 1.0, and the Δ and δ limbs can be recalculated accordingly. For the worst-case analysis, the probabilities of failure paths 1, 3, and 5 would be increased by a factor of 100, yielding a new total of $\approx .041$ for the upper bound. For the best-case analysis, success paths 1, 3, and 5 would be eliminated, and success paths 2, 4, and 6 would be increased by a factor of 10,000, yielding a new lower bound of .003.

Depending upon the situation, the nominal values of the HEPs may be used instead of their upper or lower bounds, and any other assumptions may be made regarding the probabilities of the events. Thus, the bounding analysis can be performed at various levels, each of which must be evaluated for credibility. The method is flexible, allowing any assumptions to be evaluated.

Clearly, there is no single "correct" method of performing a bounding analysis. It remains a matter of judgment as to which assumptions are credible and as to the range of situations to be explored. Note that we did not use all the error terms originally listed; we selected those with the largest estimated effects. Such a selection is adequate for a bounding analysis.

Bounding Analysis of Maintenance, Pump, or Component Test Activity Involving Two Valves (Example 1.4)

The failure event for which bounds are to be estimated is failure to return either valve (or both) to the correct position after maintenance or test. There are two major types of possible errors: completely forgetting any valve, and not returning any valve to its fully restored position.

The following PSFs were listed:

1. The valves are changed, testing or maintenance is done, and then the valves are to be restored. For a worst-case analysis, assume a

maintenance job since maintainers are oriented toward the details of a maintenance task and any other work (e.g., restoring valves) may be regarded as secondary. For a best-case analysis, assume a testing job and the use of operators.

2. Assume one or more operators or maintainers. For a worst-case analysis assume one maintainer; for the best case assume two operators, one of whom will check the other.
3. Written procedures are required by the plant; assume 50 steps as per instructions.
4. If the valves are in a remote location, special clothing (for a radiation environment) may be required. The requirement for special clothing would reduce the probability of forgetting the restoration task, but would increase the probability of errors of commission. For a worst-case analysis we increase the estimated probabilities of errors of commission and ignore the possible reduction in the probabilities of errors of omission. For a best-case analysis we consider both influences and select the most favorable. Thus, for the best case, if special clothing is required, we equate the probability of forgetting to initiate the restoration task with the probability of failing to carry out an oral instruction, that is, an HEP of .001 (.0005 to .005) (Table 20-18, p 20-28). This is the probability of failing to restore the first valve. For the worst case, we use the HEP for a maintainer, .05 (.01 to .1) (p 20-24).
5. The valves may be large and separated or small and grouped. For large valves, if the first valve is restored, with the assumption of ZD between steps in a written procedure (Table 20-20, p 20-29), the above HEP of .05 (.01 to .1) is assigned to the maintainer's failure to restore the second valve as well. For an operator, we assume an

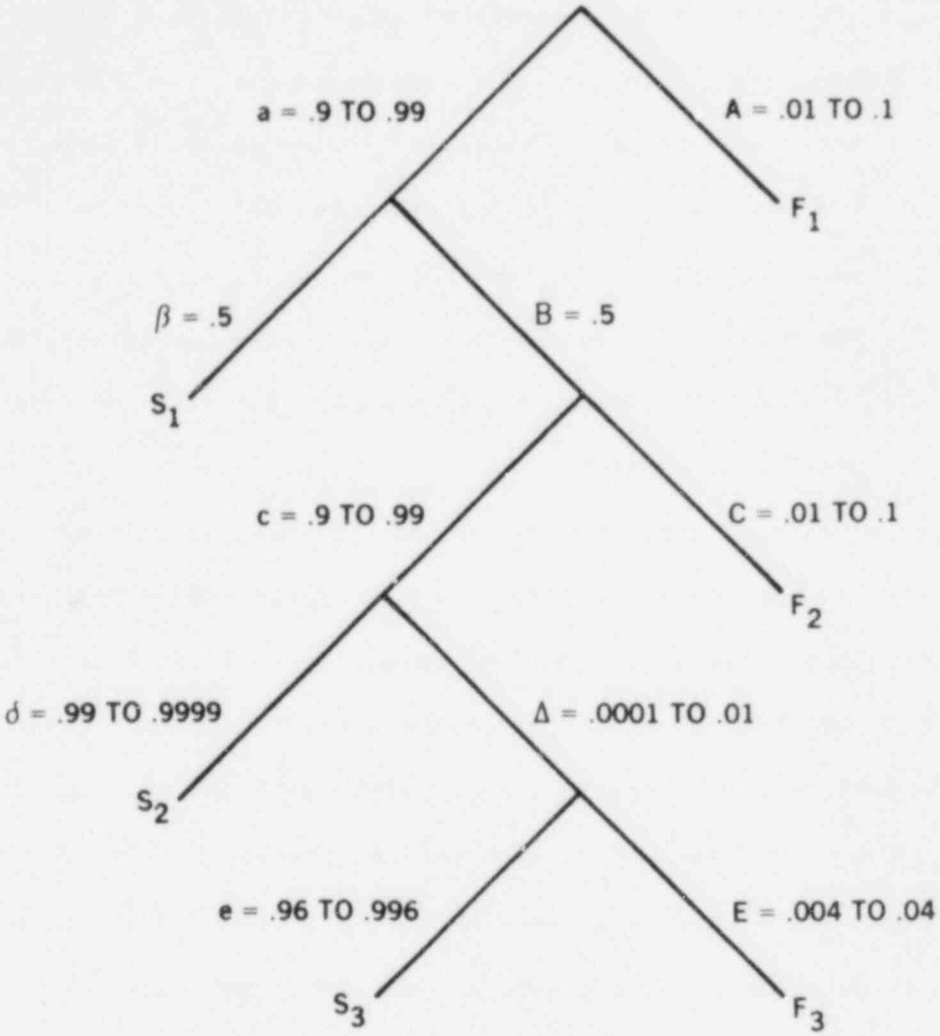
HEP of .003 (.001 to .01) for failure to restore the second valve (Table 20-20, p 20-29). For small valves that are grouped, we assume that they are operated as pairs, so that errors of omission are completely dependent. For this example, the probability of forgetting large or small valves is estimated as .5 each.

6. Probability of a sticking valve is .001 (.0001 to .01), and assume zero probability of more than one valve sticking.
7. For failure to restore a sticking valve completely, HEP = .005 (.002 to .02) (from Table 20-14, p 20-21) for a rising stem valve with no position indicator.

We will begin the bounding analysis by drawing the event tree for maintainers. In both cases (maintainers and operators) we will assume that special clothing is required and that errors of commission are increased by a factor of two when special clothing is worn. Thus, the probability of failure to restore a sticking valve completely will be taken as .01 (.004 to .04). The event tree begins with the requirement to restore the valves after completion of the maintenance or test. As stated above, the HEP for this error of omission is .001 (.0005 to .005) for operators. For maintainers this HEP is .05 (.01 to .1), as we are not allowing credit for the alerting effect of special clothing.

In Figure 21-4, note that the first success path terminates at the end of branch β. Two assumptions are made:

1. If the valves are small and grouped, there will be CD between them, so that if the person restores one he will always restore the other; therefore, branch "C" is not required.
2. If the valves are small and grouped, if one of them was sticking it would be immediately obvious, and recovery would have a probability of 1.0; therefore, branches "Δ" and "E" were not appended.



EVENTS	HEPs (BOUNDS)
A = FAILURE TO INITIATE RESTORATION	.05 (.01 TO .1)
B = VALVES ARE LARGE AND SEPARATED	.5
C = FAILURE TO RESTORE SECOND VALVE	.05 (.01 TO .1)
Δ = VALVE STICKS	.001 (.0001 TO .01)
E = FAILURE TO RESTORE A STICKING VALVE COMPLETELY, WEARING SPECIAL CLOTHING	.01 (.004 TO .04)

Figure 21-4. Event Tree for Maintainers, Example 1.4

These assumptions also apply to the event tree for operators (Figure 21-5). To calculate the worst case, we will proceed as previously, using its most pessimistic values of the HEP error bounds for maintainers. The failure paths for Figure 21-4 are:

F_1. A = .1

F_2. a-B-C = .9 x .5 x .1 = .045

F_3. a-B-c-Δ-E = .9 x .5 x .9 x .01 x .04 = .000162

.145162

Upper failure bound ≈ .15.

The event tree for an operator (which we are using for the best-case analysis) is similar to the tree for maintainers in Figure 21-4, with the addition of a checking operation by a second operator, and different values for event "C." For this analysis, we use the following success paths from Figure 21-5:

S_1. a-β = .9995 x .5 = .49975

S_2. a-B-c-δ = .9995 x .5 x .999 x .9999 = .4992003

S_3. a-B-c-Δ-e = .9995 x .5 x .999 x .0001 x .996 = .0000497

S_4. a-B-c-Δ-E-e' = .9995 x .5 x .999 x .0001 x .004 x .92 = .0000002

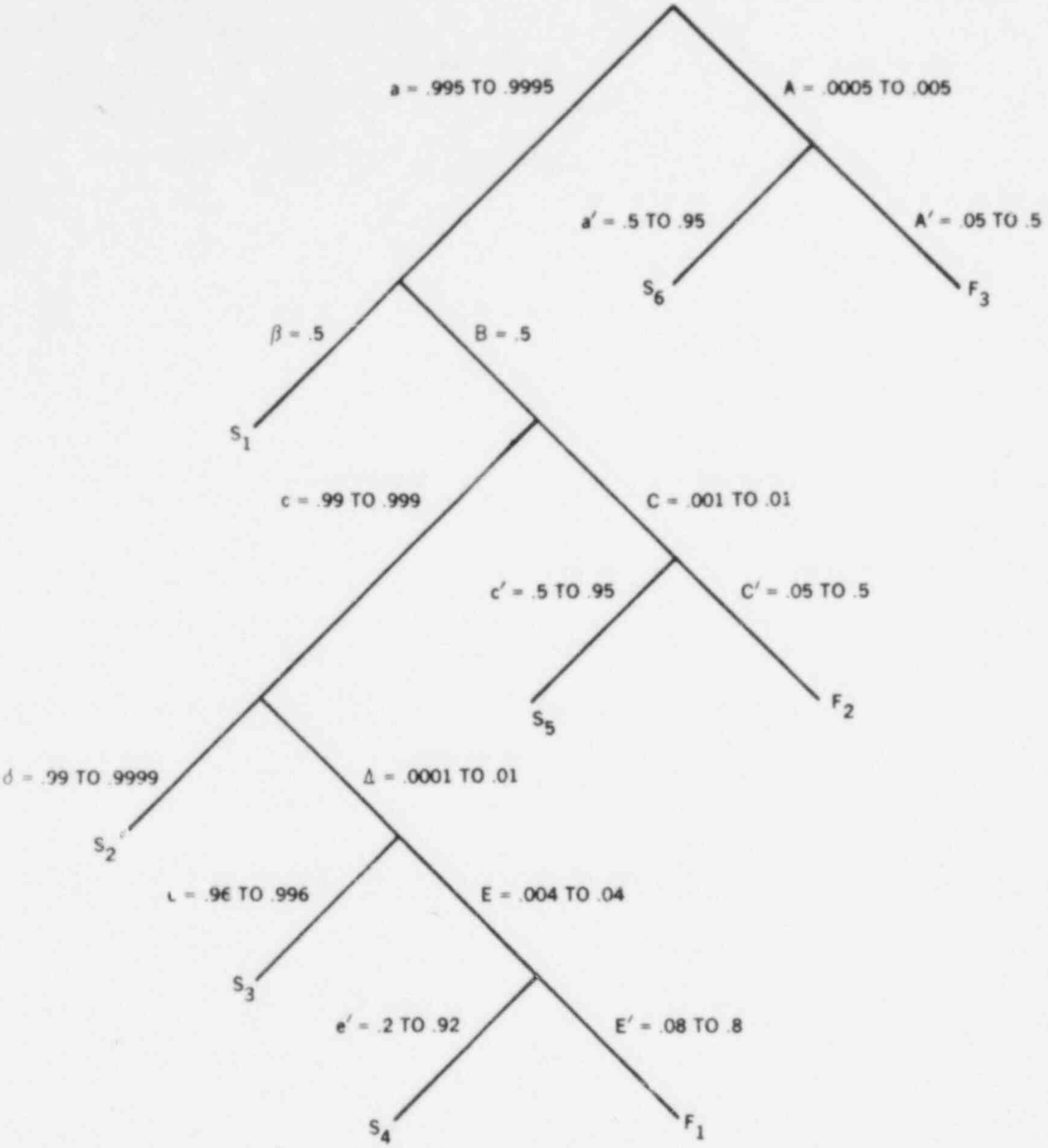
S_5. a-B-C-c' = .9995 x .5 x .001 x .95 = .0004748

S_6. A-a' = .0005 x .95 = .000475

.99995

Lower failure bound ≈ .00005.

As discussed in Example 1.3, the assumptions are flexible. The analyst may know that the proportion of valves that are small and large is other than half and half, or he may wish to assume the nominal values of the HEPs in calculating the best case. If we recalculate the best case for Figure 21-5, using the nominal values for all events in the success paths, the total



EVENTS	HEPs (BOUNDS)	SOURCE (TABLE #)
A = FAILURE TO RESTORE FIRST VALVE	.001 (.0005 TO .005)	20 - 18 (P20 - 28)
A′ = FAILURE OF CHECKER TO DETECT ERROR "A"	.1 (.05 TO .5)	20 - 16 (P20 - 25)
B = VALVES ARE LARGE AND SEPARATED	.5	
C = FAILURE TO RESTORE SECOND VALVE	.003 (.001 TO .01)	20 - 18 (P20 - 28)
C′ = FAILURE OF CHECKER TO DETECT ERROR "C"	.1 (.05 TO .5)	20 - 16 (P20 - 25)
Δ = VALVE STICKS	.001 (.0001 TO .01)	
E = FAILURE TO RESTORE STICKING VALVE COMPLETELY (HEPs DOUBLED)	.01 (.004 TO .04)	20 - 14 (P20 - 22)
E′ = FAILURE OF CHECKER TO DETECT ERROR "E"	.2 (.08 TO .8)	20 - 16 (P20 - 25)

Figure 21-5. Event Tree for Operators, Example 1.4

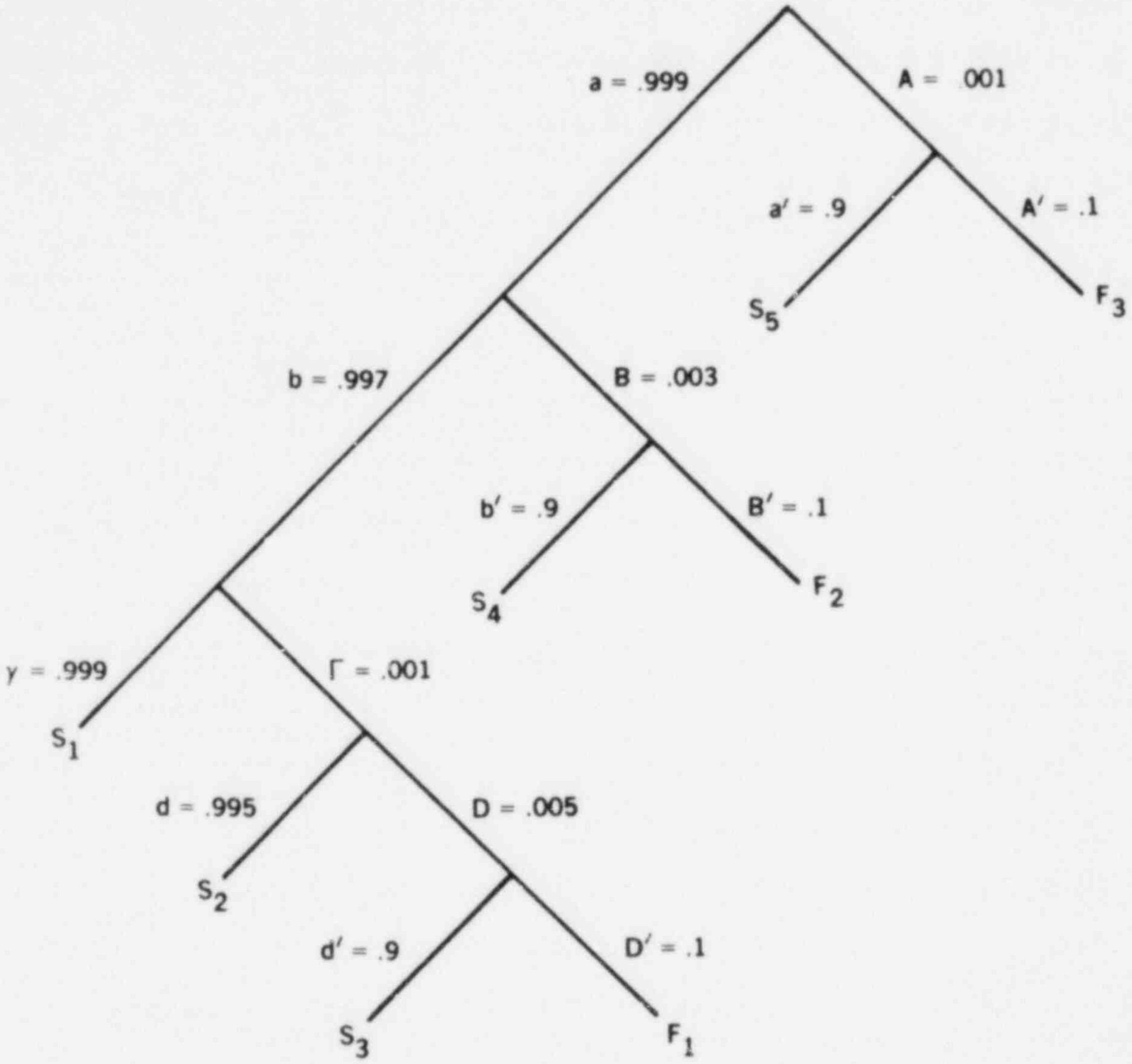
success probability will be .9997492, yielding a lower bound of .00025 which may be more credible than the first figure. Similarly, the worst-case bounds can be recalculated using whichever assumptions seem reasonable.

Bounding Analysis of Valve Line-Up Activity (Example 1.5)

The event for which bounds are to be estimated is the failure to return both of two manual valves to their correct positions after maintenance or rotation of components. As in the other problems, there are two major types of possible errors: forgetting any valve completely, and not returning any valve to its fully restored position. We will assume that special clothing is not required, and use the basic HEPs for errors of commission.

According to the instructions for item 1.5 (p 21-27), there are two possibilities: the activities are either scheduled or unscheduled.

For the best-case analysis we assume a scheduled activity, performed by an auxiliary operator, using tags, and checked by a second operator. With these assumptions the probability of failure becomes negligibly small, regardless of valve size and location. The scheduled activity is regarded as an oral instruction, with an HEP of .001 (.0005 to .005) for failure to initiate the task. With only two valves, the HEP for failure to restore the second valve is .003 (.001 to .01) (Table 20-20, p 20-29). The probability of failing to restore a sticking valve is also negligibly small, as established in Example 1.4. Figure 21-6 is the event tree for operators. Even using the nominal values of the HEPs (instead of the most optimistic values), the summed probabilities for the success paths are (see p 21-41):



EVENTS	HEPs (BOUNDS)
A = FAILURE TO RESTORE FIRST VALVE	.001 (.0005 TO .005)
A′ = FAILURE OF CHECKER TO DETECT ERROR "A"	.1 (.05 TO .5)
B = FAILURE TO RESTORE SECOND VALVE	.003 (.001 TO .01)
B′ = FAILURE OF CHECKER TO DETECT ERROR "B"	.1 (.05 TO .5)
Γ = VALVE STICKS	.001 (.0001 TO .01)
D = FAILURE TO RESTORE A STICKING VALVE COMPLETELY	.005 (.002 TO .02)
D′ = FAILURE OF CHECKER TO DETECT ERROR "D"	.1 (.04 TO .4)

Figure 21-6. Event Tree for Operators, Example 1.5

S_1. a-b-γ = .999 x .997 x .999 = .995007

S_2. a-b-Γ-d = .999 x .997 x .001 x .995 = .000991

S_3. a-b-Γ-D-d' = .999 x .997 x .001 x .005 x .9 = .0000045

S_4. a-B-b' = .999 x .003 x .9 = .0026973

S_5. A-a' = .001 x .9 = .0009

.9995998

Lower failure bound ≈ .0004.

For the worst-case analysis we assume unscheduled activities, that there has been an oral directive to do the maintenance job, that there are no written procedures, and that the restoration itself is part of the overall job but is not specifically called out. The maintenance supervisor may ask, "Did you complete the job?" but he will not ask specifically about the two valves. Under these circumstances, we assume a single maintainer (no checking) and assign the HEP of .05 (.01 to .1) from p 20-24 to either valve for failure to restore the valve after work in finished. The worst-case analysis would yield a probability of .9 x .9 = .81 ≈ .8 that both valves will be restored, which is the upper bound for this problem. The probability of a sticking valve has only a minor effect and will be ignored.

Final Comment on Bounding Analysis

The level of bounding to be used is up to the judgment of the analyst. Clearly, the bounds can be manipulated to yield results that are unrealistically optimistic or pessimistic. A realistic analysis must consider all the relevant factors and must evaluate the credibility of the various sequences of events in the probability diagrams. If only extreme values are used, almost any worst-case analysis will yield unreasonably low success estimates for the typical circumstances in NPPs, and almost any best-case analysis will result in the unrealistic conclusion that all human errors will be completely recovered.

There are no rules for the level of bounding; the appropriate levels must be determined on the basis of the application of the analysis and the requirements of the ultimate users.

CHAPTER 22. CONCLUDING COMMENTS

This chapter briefly assesses the state of technology in human reliability analysis for risk assessment and for design and development work. The possible uses of and precautions in using the handbook are described.

What the Handbook Is

This handbook presents a human reliability method, THERP, and performance models and estimated HEPs for estimating the probabilities and effects of human errors on reliability and safety in large systems like NPPs. THERP is a conventional reliability analysis method that uses event trees. The models and HEPs are based on experience, theory, and limited data.

The handbook represents a heuristic approach--admittedly imperfect but intended to stimulate further thinking and investigation. Since many of the values listed in the handbook must be regarded as hypotheses, users are urged to test the models and HEPs empirically. Their findings should be reported in the open literature, to the authors, or to the Division of Systems and Reliability Research, Office of Nuclear Regulatory Research, U.S. NRC, so that the handbook can be updated over time.

Need for the Handbook

We prepared this handbook at the request of the NRC because no such tool for the performance of in-depth human reliability analyses existed. The Rogovin Report states, "The best way to improve the existing design process [in NPPs] is by relying in a major way upon quantitative risk analyses, and by emphasizing those accident sequences that contribute significantly to risk. The design review can then focus on those plant systems that contribute to risk, identify weak points, and upgrade various requirements (maintenance, for example) to eliminate them" (Rogovin and Frampton, 1980, p 150).

Ideally, a handbook such as this should be based on data gathered from actual operating experience in NPPs. Although the need for such a data base is well-recognized, the data have not been gathered. Because of the pressing need for human reliability analyses in NPPs, we prepared the handbook as a working guide until data based on NPP experience becomes available. There are gaps in the coverage of NPP tasks, and even experienced analysts will have uncertainties about predicting performance under all the conditions that could develop in NPPs. We hope to fill such gaps as more experience and data are gained.

This handbook addresses real-world problems. To be useful in human reliability analysis, certain assumptions had to be made about the models and estimated HEPs, and some corners had to be cut. We don't believe that the shortcomings of this handbook will offend those with practical problems to solve, since our most important objective is to introduce the user to the methodology of quantitative and qualitative evaluation of human reliability. At this point, sound methodology is more important than the degree of accuracy attainable.

We are familiar with academic models of human behavior, beginning with the work of Hull (1943 and 1950) and including more recent efforts in human reliability (summarized by Meister, 1971). We find that these models are not very useful. It is still a standard joke in university psychology departments that the theories and models can predict anything after the results are in. Regrettably, most models of behavior substitute "postdiction" for prediction, or are addressed to problems of very narrow scope.

Our approach is different. The performance models and estimated HEPs are unusual in the field of behavior technology in that they can be used to predict probabilities of human errors for identifiable real-world tasks and in that these predictions are verifiable. However, there are limitations.

Limitations in Use of the Handbook

The limitations of this handbook were described in Chapter 1. At this time we reemphasize two of those limitations: the dearth of "hard" data, and the possible misuse of the handbook.

The most obvious limitation is that the HEPs have wide bounds of uncertainty. In some cases, the HEPs might be off by as much as a factor of 10 either way. There is another standard joke in psychology departments that "correct within an order of magnitude" means "wrong." Yet, for some applications, even this degree of latitude is tolerable. One reason the uncertainty bounds are wide is that they are generic, not plant-specific. When the models and HEPs are applied to a specific plant where the analyst can identify and evaluate the relevant PSFs, the uncertainty bounds can be narrowed.

Finally, the uncertainty bounds are wide because the nominal HEPs are based to a large extent on our experience (aided by psychological theory) and what little data could be found that related to NPP tasks. Thus, the major limitation to accuracy in human reliability analysis continues to be the paucity of objective probability data collected in applied settings, and in realistic simulations of these settings. Previous reviews of human reliability analysis have indicated that there is no scarcity of models, but that data are hard to find (Swain, 1964b, 1969b, and 19 7b; and Meister, 1971 and 1978; and Embrey, 1976).

A second limitation, or risk, is that the handbook may easily be misused as a "cookbook" by inexperienced "chefs." In Chapter 1 we cited Jens Rasmussen's concern about a ship's captain trying to perform surgery using a handbook. It is inevitable that the handbook will be misused by some, or used by others to "justify" inadequate designs or procedures. However, any text or handbook requiring judgment and interpretation can be misused. There is no easy remedy for this.

If the reader wishes to use the handbook in an evaluation but feels that his preparation is inadequate, there are a number of short courses offered in human factors which will introduce him to the field and help him to develop an awareness of the methods used by specialists. A most valuable benefit derived from these courses is the opportunity to meet people experienced in the field, whom one may contact for help with problems.

When to Use the Handbook

As indicated in Part II, human reliability analysis should be used in all stages of a large system, from inception to retirement. Early in the design stages of a new system, one will not have specific human performance requirements, and very wide uncertainty bounds must be used. However, even in the earliest design stages, the handbook can be used to help identify any human factors weaknesses and to provide gross estimates of risk in alternative design concepts; e.g., control room layout or automatic versus manual modes of operation. This approach has been used in military systems and NASA programs and is strongly recommended for the nuclear industry.

When a plant is operating, it is much easier to obtain accurate information on plant-specific PSFs. Often, a new design is based on an existing plant. The personnel at that plant can provide valuable suggestions regarding the overall design, plant policies, practices, procedures, training, etc, that can improve the human reliability of the new plant. In such a situation, the handbook can serve as a guide to asking the right questions and to evaluating the relative importance of suggestions received.

The uses of the handbook in evaluation of existing facilities are obvious. We hope that utility management personnel will use it to conduct in-house evaluations to aid in determining the adequacy of their established policies, procedures, and designs.

Although the handbook was written primarily for the nuclear power industry, it applies to many other industries as well. For most industrial cases, the material can be applied with only minor changes, if any.

Needs for Further Research to Improve HEP Estimates

From the preceding it is apparent that research is needed to collect both objective and subjective human performance data that are applicable to NPP tasks. Some years ago a plan for developing a human performance data bank was developed for the U.S. Navy, but, despite some initial enthusiasm, the plan was not implemented (Swain, 1971). Later the elements of this plan were repeated in the context of NPP operations (Swain, 1975). More recently, the NRC has sponsored some studies of the Licensee Event Reporting (LER) system (Joos et al, 1979; and others in progress) and is taking other steps to develop a useful system of collecting human performance data. A Paris-based organization, the Committee on the Safety of Nuclear Installations, Nuclear Energy Agency, Organization for Economic Co-Operation and Development, sponsors a "Group of Experts on Human Error Data and Assessment" (Chairman, Jens Rasmussen, Risø National Laboratory, Denmark) that includes U.S. representatives. This group is seeking to develop consensus methods for collecting estimates of human performance probabilities.

Although the status of human performance data collection should improve in the future, there are problems at present. In addition to the need for controlled experiments to answer specific questions about behavior dynamics, there is a need for data based on the experiences of NPP personnel. The present LER system does not provide such data. The LERs do provide valuable information about errors that are reported, but rarely are the important performance shaping factors related to these errors reported or described in sufficient detail for complete analysis. Also, many errors are not reported at all--these include, but are not limited to, the ones that did not result in events designated by the NRC as reportable. Thus, the numerator information needed for HEPs is inadequate (HEP = errors made ÷ opportunities for error). The denominator information is totally lacking, although in some cases it can be approximated from knowledge of schedules or interviews of personnel.

When abnormal occurrences have serious consquences (e.g., the Brown's Ferry Accident or the TMI Accident), plant personnel should be interviewed by people skilled in interviewing as well as in human factors. This was not done in either of the above incidents until long after memories began to fade or were influenced by other considerations. Without proper interviews by people qualified in the human factors area, valuable information is likely to be lost. If we are to predict how NPP personnel will perform under stress, it is essential to obtain this information as soon as possible after an abnormal event. Without such data to modify the HEPs collected in simulations of abnormal events, the use of simulator data in reliability analysis will continue to be suspect.

For some tasks, the best available source of objective data is the dynamic simulator. To date this source has been neglected in collecting HEPs for routine or dynamic tasks. The potential use of simulators for such purposes has been recognized in the last few years, and studies are being undertaken for this purpose by the Electric Power Research Institute (EPRI).

Studies that simulate tasks performed outside the control room are also needed. WASH-1400 risk assessments indicated that most of the human error impact on the availability of ESFs arose from maintenance and calibration tasks and the errors associated with restoring ESFs to their normal operating states, rather than from control room activities. This area of research has not yet been addressed.

Clearly, it will take time to collect the large body of human performance data needed. In the meantime, the NRC is investigating an interim solution that involves the use of expert judgment. In his article on the use of ordinal scaling for human reliability analysis, Rook (1964) noted that, although people are not good estimators of absolute error probabilities, they can reliably rank-order human tasks in terms of some single dimension such as task difficulty, error-likeliness, or danger. Psychological scaling, as this method is called, is not new. The basic techniques have been in use since the 1920s. A more recently developed technique, called the Delphi Technique (Dalkey and Helmer, 1963), is an attempt to have experts directly estimate the failure probabilities of various events. There are serious drawbacks to this technique (Sackman, 1975, and Swain, 1977b). For example, in one application (Shooman and Sinkar, 1977) 12 subject-matter experts made estimates of the probabilities of 10 hazardous events. The difference between the lowest and highest estimates for each of the 10 events ranged from a low of a factor-of-200 difference on up to a factor-of-50000 difference--clearly demonstrating that this technique is not a highly reliable scaling device. Despite such problems, the value of having experts directly assign probabilities of failure to system events (including human errors) is obvious. The NRC is sponsoring research to see if the method's reliability can be improved. If the variability among judges can be reduced to a reasonable level, the approach should be useful at least for bounding analyses and for comparative evaluations of different design concepts. For those skeptical of using subjective data in risk assessments, Meister (1978, p 383), in his article on the use of subjective data in human reliability estimates, makes two relevant points:

> "(1) One should not think of subjective and objective data as an irreconcilable dichotomy. The subjective ... data bank will complement the objective one and may be able to solve design problems for which presently available data are not suitable. (2) Efforts to validate and revise the subjective ... data bank should proceed concurrently with its use. Ultimately the subjective data bank may be transformed into an acceptably objective form. Until that time comes, it will help us to do what needs to be done."

What is needed, then, are objective and subjective data banks of human performance. The objective data bank will consist of HEPs observed "on the job" and the PSFs associated with the tasks. The subjective data bank will consist of expert opinions quantified by psychological scaling techniques.

Each of the above research areas can provide data that relate errors to different ergonomics features of equipment and to different types of procedures, thereby reducing the uncertainty of our estimates. Some people have said that ergonomics changes are frosting on the cake and represent an unaffordable luxury. We hope that no one will use our imperfect data to justify less than the best available ergonomics in a future plant, or to justify failure to implement reasonable ergonomic improvements to existing plants. Our human performance models are not sufficiently fine-tuned to quantify the reduction in error to be derived from incorporating every recommended principle of ergonomics. But we know from experience that a system which does not incorporate standard ergonomics practices is a system at risk.

REFERENCES

1. Ablitt, J. F., A Quantitative Approach to the Evaluation of the Safety Function of Operators of Nuclear Reactors, AHSB(S)R-160, Authority Health and Safety Branch, United Kingdom Atomic Energy Authority, Risley, England, 1969.

2. Appley, M. H. and Trumbull, R. (eds), Psychological Stress, New York: Appleton-Century-Crofts, 1967.

3. Attneave, F., Applications of Information Theory to Psychology, New York: Holt-Dryden, 1959.

4. Axelsson, R. and Lundberg, U., Working Environment of Operating Personnel in a Nuclear Power Plant: Pilot Study, Report No. TA 875-R1, Professor Sten Luthander Engineering Office, Bromma, Sweden, December 1975.

5. Baker, C. A. and Grether, W. F., Visual Presentation of Information, WADC TR54-160, Wright Air Development Center, Wright-Patterson Air Force Base, OH, 1954.

6. Barlow, R. E. and Proschan, F., Mathematical Theory of Reliability, New York: Wiley, 1965.

7. Bartlett, F. C., "The Bearing of Experimental Psychology Upon Human Skilled Performance," British Journal of Industrial Psychology, 1951, 8, 209-217.

8. Bell, B. J., A Review of The Literature on Psychological Stress, Sandia National Laboratories, Albuquerque, NM, 1980, in preparation.

9. Berelson, B. and Steiner, G. A., Human Behavior: An Inventory of Scientific Findings, New York: Harcourt, Brace, 1964.

10. Berkun, M. M., "Performance Decrement Under Psychological Stress," Human Factors, 1964, 6, 21-30.

11. Berkun, M. M., Bialek, H. M., Kern, R. P., and Yagi, K., "Experimental Studies of Psychological Stress in Man," Psychological Monographs: General and Applied, 1962, 76, Whole No.

12. Blanchard, R. E., Mitchell, M. B., and Smith, R. L., Likelihood of Accomplishment Scale for a Sample of Man-Machine Activities, Dunlap & Assoc., Inc., Santa Monica, CA, 1966.

13. Bloom, W., Ch. 53, "Shift Work and Human Efficiency," in Fleishman, E. A. (ed), Studies in Personnel and Industrial Psychology, Homewood, IL: Dorsey Press, 1961.

14. Bond, N. A., Jr., "Some Persistent Myths About Military Electronics Maintenance," Human Factors, 1970, 12, 241-252.

15. Brooks, F. A., Jr., "Operational Sequence Diagrams," IRE Transactions on Human Factors in Electronics, Vol. HFE-1, No. 1, March 1960, pp 33-34.

16. Brune, R. L. and Weinstein, M., Development of a Checklist for Evaluating Maintenance, Test, and Calibration Procedures Used in Nuclear Power Plants, NUREG/CR-1368, SAND80-7053, HPT Inc., Thousand Oaks, CA, 1980.

17. Brune, R. L. and Weinstein, M., Procedures Evaluation Checklist for Maintenance, Test, and Calibration Procedures, NUREG/CR-1369, SAND80-7054, HPT, Inc., Thousand Oaks, CA, 1980.

18. Bryan, G. L., Bond, N. A., LaPorte, H. R., and Hoffman, L. S., Electronics Troubleshooting: A Behavorial Analysis, U.S.C Electronics Personnel Research Group Report No. 13, University of Southern California, Los Angeles, CA, 1956.

19. Buckner, D. N., and McGrath, J. J., (eds), Vigilance: A Symposium, New York: McGraw Hill, 1963.

20. Chambers, A. N., Development of a Taxonomy of Human Performance: A Heuristic Model for the Development of Classification Systems, AIR-726-3/69-TR-4A, American Institutes for Research, Washington, DC, August 1969.

21. Chapanis, A., Research Techniques in Human Engineering, Baltimore: Johns Hopkins Press, 1959.

22. Chapanis, A., "Men, Machines, and Models," American Psychologist, 1961, 16, 113-131.

23. Chapanis, A., Man-Machine Engineering, Belmont, CA: Wadsworth Publishing Co., 1965.

24. Chapanis, A., Garner, W. R., and Morgan, C. T., Applied Experimental Psychology, New York: Wiley, 1949.

25. Chapanis, A. and Scarpa, L. C., "Readability of Dials at Different Distances with Constant Visual Angle," Human Factors, 1967, 9, 415-426.

26. Colquhoun, W. P., "Circadian Rhythms, Mental Efficiency, and Shift Work," Ergonomics, 1970, 13, 558-560.

27. Colquhoun, W. P., Blake, M. J. F., and Edwards, R. S., "Experimental Studies of Shift-Work I: A Comparison of 'Rotating' and 'Stabilized' 4-Hour Shift Systems," Ergonomics, 1968, 11, 437-453.

28. Colquhoun, W. P., Blake, M. J. F., and Edwards, R. S., "Experimental Studies of Shift-Work II: Stabilized 8-Hour Shift Systems," Ergonomics, 1968, 11, 527-546.

29. Colquhoun, W. P., Blake, M. J. F., and Edwards, R. S., "Experimental Studies of Shift-Work III: Stabilized 12-Hour Shift Systems," Ergonomics, 1969, 12, 865-882.

30. Coombs, C. H., The Theory of Psychological Scaling, Univ. of Michigan Eng. Bull. No. 34, Ann Arbor, MI, 1952.

31. Craig, A. and Colquhoun, W. P., "Vigilance: A Review," in C. G. Drury and J. G. Fox (eds), Human Reliability in Quality Control, London: Taylor and Francis, 1975, 71-88.

32. Dalkey, N. and Helmer, F., "An Experimental Application of the DELPHI Method to the Use of Experts," Management Sciences, 1963, 9, 458-467.

33. Eaton, D. L., "NIE Attacks the Reading and Language Problem," American Federation, May 1974, 35-36.

34. Edwards, A. L., Techniques of Attitude Scale Construction, New York: Appleton-Century-Crofts, 1957.

35. Edwards, E. and Lees, F. P., Man and Computer in Process Control, London: The Institution of Chemical Engineers, 1973.

36. Edwards, E. and Lees, F. P., (eds), The Human Operator in Process Control, London: Taylor & Francis, 1974.

37. Embrey, D. E., Human Reliability in Complex Systems: An Overview, NCSR.R10, National Centre of Systems Reliability, United Kingdom Atomic Energy Authority, Warrington, England, July 1976.

38. English, H. B. and English, A. C., A Comprehensive Dictionary of Psychological and Psychoanalytical Terms, New York: David McKay Co., 1958.

39. EPRI Review - See Seminara et al, 1976.

40. Evans, R. A., "Reliability Optimization," in E. J. Henley and J. W. Lynn (eds), Generic Techniques in Systems Reliability Assessment, Leyden, The Netherlands: Noordhoff International Publishing, 1976, 117-131.

41. Federal Standard No. 595A, Colors, U.S. General Services Administration, Jan. 2, 1968 (Reprinted March 1974).

42. Fitts, P. M. (ed), Human Engineering For An Effective Air Navigation and Traffic Control System, National Research Council, Washington DC, 1951.

43. Fitts, P. M., and Posner, M. I., Human Performance, Belmont, CA: Brooks/Cole Publishing Co., 1967.

44. Fivars, G., and Gosnell, D., Nursing Evaluation: The Problem and the Process, New York: Macmillan Co., 1966.

45. Flanagan, J. C., "The Critical Incident Technique," Psychol. Bull., 1954, 51, 327-358.

46. Fleishman, E. A., Kincade, R. G., and Chambers, A. N., Development of a Taxonomy of Human Performance: A Review of the First Year's Progress, AIR-726-11/68-TPRI, American Institutes for Research, Washington, DC, November 1968.

47. Fleishman, E. A., Teichner, W. H., and Stephenson, R. W., Development of a Taxonomy of Human Performance: A Review of the Second Year's Progress, AIR-726-1/70-TPR2, American Institutes for Research, Washington, DC, January 1970.

48. Folkard, S., Monk, T. H., and Lobban, M. C., "Towards a Predictive Test of Adjustment to Shift Work," Ergonomics, 1979, 22, 79-91.

49. Fox, J. G., "Vigilance and Arousal: A Key to Maintaining Inspectors' Performance," in C. G. Drury and J. G. Fox (eds), Human Reliability in Quality Control, London: Taylor and Francis, 1975, 89-96.

50. Frost, G., "Man-Machine Dynamics," Ch. 6 in H. P. VanCott and R. G. Kinkade (eds), Human Engineering Guide to Equipment Design, (rev ed), Washington, DC: Government Printing Office, 1972.

51. Gabriel, R. F., and Burrows, A. A., "Improving Time-Sharing Performance of Pilots Through Training," Human Factors, 1968, 10, 33-40.

52. Garner, W. R., Uncertainty and Structure as Psychological Concepts, New York: Wiley, 1962.

53. Grandjean, E., "Fatigue: Its Physiological and Psychological Significance," Ergonomics, 1968, 11, 427-436.

54. Grandjean, E. P., Wotzka, G., Schaad, R., and Gilgen, A., "Fatigue and Stress in Air Traffic Controllers," Ergonomics, 1971, 14, 159-165.

55. Grant, J. S., "Concepts of Fatigue and Vigilance in Relation to Railway Operation," Ergonomics, 1971, 14, 111-118.

56. Green, A. E., Safety Assessment of Automatic and Manual Protective Systems for Reactors, AHSB(S)R-172, Authority Health and Safety Branch, United Kingdom Atomic Energy Authority, Risley, England, 1969.

57. Green, A. E., and Bourne, A. J., Reliability Technology, London: Wiley-Interscience, 1972.

58. Grinker, R. R., and Spiegel, J. P., Men Under Stress, New York: McGraw-Hill, 1963. (Reprinted from 1945)

59. Guilford, J. P., Psychometric Methods, New York: McGraw-Hill, 1954.

60. Gulliksen, H. and Messick, S. (eds), Psychological Scaling: Theory and Applications, New York: Wiley, 1960.

61. Guttmann, H. E. and Finley, B. H., "Accuracy of Visual Spatial Interpolation," Ergonomics, 1970, 13, 243-246.

62. Haney, R. W., "The Effects of Instructional Format on Functional Testing Performance," Human Factors, 1969, 11, 181-187.

63. Harris, D. H. and Chaney, F. B., Human Factors in Quality Assurance, Publ. P7-2787/501, Autonetics Div. of North American Rockwell Corp., Anaheim, CA, 1967.

64. Harris, D. H. and Chaney, F. B., Human Factors in Quality Assurance, New York: Wiley, 1969.

65. Harris, W., Mackie, R. R., and Wilson, C. L., Performance Under Stress: A Review and Critique of Recent Studies. Human Factors Research, Inc., Research on the Development of Performance Criteria Technical Report VI, Los Angeles, CA, July 1956.

66. Hertzberg, H. T. E., "Engineering Anthropology," Ch. 11 in H. P. VanCott and Kinkade, R. G. (eds), Human Engineering Guide to Equipment Design, (rev ed), Washington, DC: U.S. Government Printing Office, 1972.

67. Hilgard, E. R., Atkinson, R. C., and Atkinson, R. L., Introduction to Psychology, (5th ed), New York: Harcourt Brace Jovanovich, 1971.

68. Hull, C. L., Principles of Behavior, New York: Appleton-Century, 1943.

69. Hull, C. L., "Behavior Postulates and Corollaries - 1949," Psychol. Review, 1950, 57, 173-180.

70. "Human Reliability Analysis," Section 6.1 in Appendix III-Failure Data, of WASH-1400 (NUREG-75/014): Reactor Safety Study -- An Assessment of Accident Risks in U.S. Commercial Nuclear Power Plants, U.S. Nuclear Regulatory Commission, Washington, DC, Oct. 1975.

71. Hurst, R. (ed), Pilot Error, London: Crosby Lockwood Staples, 1976.

72. Irwin, I. A., Levitz, J. J., and Freed, A. M., Human Reliability in the Performance of Maintenance, Report LRP 317/TDR-63-218, Aerojet-General Corp., Sacramento, CA, May 1964.

73. Irwin, I. A., Levitz, J. J., and Ford, A. A., "Human Reliability in the Performance of Maintenance," Proceedings of the Symposium on quantification of Human Performance, Electronics Industries Assn. and Univ. of New Mexico, Albuquerque, NM, Aug. 1964, 143-198.

74. Jerison, H. J. and Pickett, R. M., "Vigilance: A Review and Reevaluation," Human Factors, 1963, 5, 211-238.

75. Joos, D. W., Sabri, Z. A., and Husseiny, A. A., "Analysis of Gross Error Rates in Operation of Commerical Nuclear Power Stations," Nuclear Engineering and Design, 1979, 52, 265-300.

76. Kappauf, W. E. and Smith, W. M., Design of Instrument Dials for Maximum Legibility - II, A Preliminary Experiment on Dial Size and Graduation, MCREXD-694-IN, Air Material Command, Wright-Patterson Air Force Base, OH, 1948.

77. Kemeny, J. G. (Chairman), Report of the President's Commission on the Accident at Three Mile Island, U.S. Government Printing Office, Washington, DC, October 1979.

78. Kendall, M. G., Rank Correlations Methods, London: Chas. Griffin, 1948.

79. Kletz, T. A., "The Man in the Middle," Safety Newsletter, No. 123, May 1979.

80. Kletz, T. A. and Whitaker, G. D., Human Error and Plant Operation, EDN 4099, Safety and Loss Prevention Group, Petrochemicals Division, Imperial Chemical Industries, Ltd., Billingham, England, 1973.

81. Klier, S. and Linskey, J. W., Selected Abstracts from the Literature on Stress, NAVTRADEVCEN 565-1, Port Wash., NY, U.S. Naval Training Device Center, November 1960.

82. Kurke, M. I., "Evaluation of a Display Incorporating Quantitative and Check Reading Characteristics," J. Appl. Psychol., 1956, 40, 233.

83. Kurke, M. I., "Operational Sequence Diagrams in System Design," Human Factors, 1961, 3, 66-73.

84. Mackie, R. R. (ed), Vigilance: Theory, Operational Performance, and Physiological Correlates, Vol. 3 Human Factors, NATO Conference Series, New York: Plenum Press, 1977.

85. Malone, T. B., Kirkpatrick, M., Mallory, K., Eike, D., Johnson, J. H., and Walker, R. W., Human Factors Evaluation of Control Room Design and Operator Performance at Three Mile Island-2, Vols. I, II, III, NUREG/CR-1270, Three Mile Island Special Inquiry Group, U.S. Nuclear Regulatory Commission, Washington, DC: January 1980.

86. Marshall, S. L. A., Men Against Fire, Gloucester, MA: Peter Smith, 1978.

87. Matarazzo, J. D., Wechsler's Measurement and Appraisal of Adult Intelligence, (5th ed), Baltimore: Williams & Wilkins, 1972.

88. Maurice, M., Shift Work, Geneva: International Labour Office, 1975.

89. McConnell, J. V., "Biorhythms: A Report and Analysis," Journal of Biological Psychology, 1978, 20, 13-24.

90. McCormick, E. J., Human Factors in Engineering and Design, 4th ed, New York: McGraw-Hill, 1975.

91. McCornack, R. L., Inspector Accuracy: A Study of the Literature, SCTM-53-61(14), Sandia National Laboratories, Albuquerque, NM, 1961.

92. McFarland, R. A., "Understanding Fatigue in Modern Life," Ergonomics, 1971, 14, 1-10.

93. McGeoch, J. A., The Psychology of Human Learning, New York: Longmans, Green, 1942.

94. McGeoch, J. A. and Irion, A. L., The Psychology of Human Learning (2nd ed), New York: Longmans, Green, 1952.

95. McKenzie, R. M., "On the Accuracy of Inspectors," Ergonomics, 1958, 1, 258-272.

96. Meister, D. "The Problem of Human-Initiated Failures," Proceedings of 8th National Symposium on Reliability and Quality Control, Inst. of Electrical and Electronics Engineers, New York, 234-239, January 1962.

97. Meister, D., "Methods of Predicting Human Reliability in Man-Machine Systems," Human Factors, 1964, 6, 621-646.

98. Meister, D., "Human Factors in Reliability," Section 12 in W. G. Ireson (ed), Reliability Handbook, New York: McGraw-Hill, 1966.

99. Meister, D., "Tables for Predicting the Operational Performance of Personnel," Appendix A in J. J. Hornyak, Effectiveness of Display Subsystem Measurement and Prediction Techniques, RADC-TR-67-292, Rome Air Development Center, Griffiss Air Force Base, NY, September 1967.

100. Meister, D., Comparative Analysis of Human Reliability Models, Bunker-Ramo Electronics Systems Division, Westlake Village, CA, Report L0074-107, November 1971.

101. Meister, D., "Subjective Data in Human Reliability Estimates," Proceedings 1978 Annual Reliability and Maintainability Symposium, Inst. of Elect. and Electronics Engineers, New York, 380-384, January 1978.

102. Meister, D. and Rabideau, G. F., Human Factors Evaluation In System Development, New York: Wiley, 1965.

103. MIL-H-46855B, Military Specification, Human Engineering Requirements for Military Systems, Equipment and Facilities, U.S. Dept. of Defense, Washington, DC, 31 January 1979.

104. MIL-STD-1472B, Military Standard, Human Engineering Design Criteria for Military Systems, Equipment and Facilities, U.S. Dept. of Defense, Washington, DC, 31 December 1974, with Notice 1 dated 10 May 1976 and Notice 2 dated 10 May 1978.

105. Miller, R. B., Handbook on Training and Rating Equipment Design, WADC TR 53-136, Wright Air Development Center, Wright-Patterson Air Force Base, OH, June 1953.

106. Miller, R. B., A Method for Man-Machine Task Analysis, WADC TR 53-137, Wright Air Development Center, Wright-Patterson Air Force Base, OH, June 1953.

107. Mills, R. G. and Hatfield, S. A., "Sequential Task Performance: Task Module Relationships, Reliabilities, and Times," Human Factors, 1974, 16, 117-128.

108. Mintz, A. and Blum, M. L., "A Re-examination of Nuclear Accident Proneness Concept," Ch. 57 in E. A. Fleishman (ed), Studies in Personnel and Industrial Psychology, Homewood, IL: Dorsey Press, 1961.

109. Morrissey, J. H., "New Method for the Assignment of Psychometric Scale Values from Incomplete Paired Comparisons," *J. Optical Soc. of Am.*, 1955, *45*, 373-378.

110. Müller, F. W., *Elementary Reliability Technology*, Sandia National Laboratories, Albuquerque, NM, SC-R-64-198, July 1964.

111. Munger, S. J., Smith, R. W., and Payne, D., *An Index of Electronic Equipment Operability: Data Store*, AIR-C43-1/62-RP(1), American Institutes for Research, Pittsburgh, PA, January 1962.

112. Murrell, K. F. H., Laurie, W. D., and McCarthy, C., "The Relationship Between Dial Size, Reading Distance, and Reading Accuracy," *Ergonomics*, 1958, *1*, 182-190.

113. Murrell, K. F. H., *Ergonomics: Man in His Working Environment*, London: Chapman and Hall, 1969.

114. Nielsen, D. C., *The Cause/Consequence Diagram Method as a Basis For Quantitative Accident Analysis*, Risø-M-1374, Risø National Laboratory, Roskilde, Denmark, May 1971.

115. Nielsen, D. S., *Use of Cause-Consequence Charts in Practical Systems Analysis*, Risø-M-1743, Risø National Laboratory, Roskilde Denmark, September 1974.

116. O'Hanlon, J. F. and McGrath, J. J., *Studies of Human Vigilance, an Omnibus of Technical Reports*, Human Factors Research Inc., Goleta, CA, 1968.

117. Osborne, E. E. and Vernon, H. M., *Two Contributions to the Study of Accident Causation*, IFRB Report No. 19, Her Majesty's Stationery Office, London, 1922.

118. O'Shell, H. E. and Bird, F. E., "Incident Recall," National Safety News, 1969, 100, 58-63.

119. Payne, D. and Altman, J. W., An Index of Electronic Equipment Operability: Report of Development, AIR-C-43-1/62-FR, American Institutes for Research, Pittsburgh, PA, January 1962.

120. Payne, D., Altman, J. W., and Smith, R. W., An Index of Electronic Equipment Operability: Instruction Manual, AIR-C-43-1/62-RP(3), American Institutes for Research, Pittsburgh, PA, January 1962.

121. Peter, L. J. and Hull, R., The Peter Principle, New York: William Morrow and Co., 1969.

122. Pressey, P. E., "Pilot's Review," Ch. 1 in R. Hurst (ed) Pilot Error, London: Crosby Lockwood Staples, 1976.

123. Pyrczak, F. and Roth, D., "The Readability of Directions on Non-Prescription Drugs," Journal of the American Pharmaceutical Association, 1976, 16, 242-243.

124. Rasmussen, J., "The Role of the Man-Machine Interface in Systems Reliability," E. J. Henley and J. W. Lynn, (eds) Generic Techniques in Systems Reliability Assessment, Leyden, The Netherlands: Noordhoff International Publishing, 1976, 315-323.

125. Rasmussen, J. and Jensen, A., "Mental Procedures in Real-Life Tasks: A Case Study of Electronic Trouble Shooting," Ergonomics, 1974, 17, 293-307.

126. Rasmussen, J. and Taylor, J. R., Notes on Human Factors Problems in Process Plant Reliability and Safety Prediction, Risø-M-1894, Risø National Laboratory, Roskilde, Denmark, September 1976.

127. Raudenbush, M. H., Evaluation of Operator Response at a Miniaturized Control Board of a Large Nuclear Power Reactor, Master of Science Thesis, Dept. of Nuclear Engineering, North Carolina State University at Raleigh, Raleigh, NC, 1971.

128. Raudenbush, M. H., "Human Engineering Factors in Control Board Design for Nuclear Power Plants," Nuclear Safety, 1973, 14, 21-26.

129. Remmers, H. H., Introduction to Opinion and Attitude Measurement, New York: Harper, 1954.

130. Rigby, L. V., "The Sandia Human Error Rate Bank (SHERB)," R. E. Blanchard and D. H. Harris (eds), Man-Machine Effectiveness Analysis, A Symposium of the Human Factors Society, Los Angeles Chapter, Los Angeles, CA, June 1967, 5-1 to 5-13.

131. Rigby, L. V., "The Nature of Human Error," Annual Technical Conference Transactions of the ASQC, American Society for Quality Control, Milwaukee, WI, May 1970, 45-466.

132. Rigby, L. V. and Edelman, D. A., "A Predictive Scale of Aircraft Emergencies," Human Factors, 1968, 10, 475-782.

133. Rigby, L. V. and Edelman, D. A., An Analysis of Human Variability in Mechanical Inspection, SC-RR-68-282, Sandia National Laboratories, Albuquerque, NM, May 1968.

134. Rigby, L. V. and Swain, A. D., "Effects of Assembly Error on Product Acceptability and Reliability," Proceedings of the 7th Annual Reliability and Maintainability Conference, American Society of Mechanical Engineers, New York, July, 1968, 3-12 to 3-19.

135. Rigby, L. V. and Swain, A. D., "Some Human-Factor Applications to Quality Control in a High Technology Industry," C. G. Drury and J. G. Fox (eds), Human Reliability in Quality Control, London: Taylor and Francis, 1975, 201-216.

136. Roberts, N. H., Haasl, D. F., Vesely, W. E., and Goldberg, F. F., Fault Tree Handbook, NUREG-0492, U.S. Nuclear Regulatory Commission, Washington, DC, March 1980.

137. Robbins, J. E., McKendry, J. M., and Hurst, P. M., Task-Induced Stress: A Literature Survey, AD667272, Aerospace Medical Laboratory, Wright-Patterson Air Force Base, OH, 1961.

138. Rogovin, M. and Frampton, G. E., Jr. (eds), Three Mile Island, A Report to the Commissioners and to the Public, Vol. 1, Special Inquiry Group, U.S. Nuclear Regulatory Commission, Washington, DC, January, 1980.

139. Ronan, W. W., Training for Emergency Procedures in Multi-Engine Aircraft, AIR-153-53-FR-44, American Institutes for Research, Pittsburgh, PA, March 1953.

140. Rook, L. W., Reduction of Human Error in Industrial Production, SCTM,93-62(14), Sandia National Laboratories, Albuquerque, NM, June 1962.

141. Rook, L. W., "A Method for Calculating the Human Error Contribution to System Degradation," in A. D. Swain (Chairman), Human Error Quantification, SCR-610, Sandia National Laboratories, Albuquerque, NM, April 1963.

142. Rook, L. W., Jr., "Evaluation of System Performance from Rank-Order Data," Human Factors, 1964, 6, 533-536.

143. Rook, L. W., *Motivation and Human Error*, SC-TM-65-135, Sandia National Laboratories, Albuquerque, NM, September 1965.

144. Rubinstein, E., "The Accident that Shouldn't Have Happened," *IEEE Spectrum*, 1979, *16*, 33-42.

145. Sackman, H., *Delphi Critique*, Lexington, MA: D. C. Heath & Co., 1975.

146. Schmall, T. M. (ed), *Conference Record for 1979 IEEE Standards Workshop on Human Factors and Nuclear Safety*, Inst. of Electrical and Electronic Engineers, New York, 1980.

147. Seminara, J. L., Eckert, S. K., Seidenstein, S., Gonzalez, W. R., Stempson, R. L., and Parsons, S. O., *Human Factors Methods for Control Room Design*, EPRI, NP-1118-SY, Electric Power Research Institute, Palo Alto, CA, June 1979.

148. Seminara, J. L., Gonzalez, W. R., and Parsons, S. O., *Human Factors Review of Nuclear Power Plant Control Room Design*, EPRI NP-309, Electric Power Research Institute, Palo Alto, CA, November 1976.

149. Seminara, J. L., Pack, R. W., Gonzalez, W. R., and Parsons, S. O., "Human Factors in the Nuclear Control Room," *Nuclear Safety*, 1977, *18*, 774-790.

150. Seminara, J. L., Pack, R. W., Seidenstein, S., and Eckert, S. K., "Human Factors Engineering Enhancement of Nuclear Power Plant Control Rooms," *Nuclear Safety*, 1980, *21*, 351-363.

151. Seminara, J. L., Parsons, S. O., and Parris, H. L., "Human Factors Aspects of Nuclear and Fossil Fuel Power Plant Maintainability," *Proceedings of the Human Factors Society 23rd Annual Meeting*, Santa Monica, CA, November 1979, Addendum pp 1-5.

152. Shannon, C. E. and Weaver, W., The Mathematical Theory of Communication, Urbana, IL: Univ. of Illinois Press, 1949.

153. Shapero, A., Cooper, J. I., Rappaport, J., Schaeffer, R. H., and Bates, C., Jr., Human Engineering Testing and Malfunction Data Collection in Weapon System Test Programs, WADC TR 60-36, Wright Air Development Center, Wright-Patterson Air Force Base, OH, February 1960.

154. Sheridan, T. B., "Human Error in Nuclear Power Plants," Technology Review, February 1980, 23-33.

155. Sheridan, T. B., and Ferrell, W. R., Man-Machine Systems: Information, Control, and Decision Models of Human Performance, Cambridge, MA: The MIT Press, 1974.

156. Shooman, M. L. and Sinkar, S., "Generation of Reliability and Safety Data by Analysis of Expert Opinion," Proceedings of 1977 Annual Reliability and Maintainability Symposium, Institute of Electrical and Electronic Engineers, New York, January 1977, 186-193.

157. Siegel, A. I. And Wolf, J. A., Man-Machine Simulation Models, New York: Wiley, 1969.

158. Smith, R. L. and Lucaccini, L. F., "Vigilance Research: Its Application to Industrial Problems," Human Factors, 1969, 11, 149-156.

159. Smith, R. W., and Payne, D., An Index of Electronic Equipment Operability: Sample Equipment Evaluation, AIR-C-43-1/62-RP(2), American Institutes for Research, Pittsburgh, PA, January 31, 1962.

160. Stello, V. Jr., (ed), Investigation into the March 28, 1979, Three Mile Island Accident by Office of Inspection and Enforcement, NUREG-0600, Investigate Report No. 50-320/79-10, Office of Inspection and Enforcement, U.S. Nuclear Regulatory Comm., Washington, DC, August, 1979.

161. Stevens, S. S. (ed), *Handbook of Experimental Psychology*, New York: Wiley, 1951.

162. Strong, E. K. Jr., "The Effect of Time-Interval upon Recognition Memory," *The Psychological Review*, 1913, *20*, 339-372.

163. Sugarman, R., "Nuclear Power and the Public Risk," *IEEE Spectrum*, 1979, *16*, 59-79.

164. Swain, A. D., *Maintenance Diagrams for the Preventive Maintenance of Ground Electronic Equipment: Phase I, Recommendations for Determining Their Need*, RADC TR 56-141, Rome Air Development Center, Griffiss Air Force Base, NY, November 1956.

165. Swain, A. D., *System and Task Analysis, A Major Tool for Designing the Personnel Subsystem*, SCR-457, Sandia National Laboratories, Albuquerque, NM, January 1962.

166. Swain, A. D., "Summary," in A. D. Swain (Chairman), *Human Factor Quantification*, SCR-610, Sandia National Laboratories, Albuquerque, NM, April 1963.

167. Swain, A. D., *A Method for Performing a Human Factors Reliability Analysis*, Monograph SCR-685, Sandia National Laboratories, Albuquerque, NM, August 1963.

168. Swain, A. D., "Human Factors in Design of Reliable Systems," *Proceedings of 10th National Symposium on Reliability and Quality Control*, Institute of Electrical and Electronic Engineers, New York, 250-259, January 1964.

169. Swain, A. D., "Some Problems in the Measurement of Human Performance in Man-Machine Systems," *Human Factors*, 1964, *6*, 687-700.

170. Swain, A. D., "Field Calibrated Simulation," *Proceedings of the Symposium on Human Performance Quantification in Systems Effectiveness*, Naval Material Command and the National Academy of Engineering, Washington, DC, January 1967, IV-A-1 to IV-A-21.

171. Swain, A. D., "Some Limitations in Using the Simple Multiplicative Model in Behavior Quantification," in W. B. Askren (ed), *Symposium on Reliability of Human Performance in Work*, AMRL-TR-67-88, Air National Research Laboratory, Wright-Patterson Air Force Base, OH, May 1967, 17-31.

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

173. Swain, A. D., "Overview and Status of Human Factors Reliability Analysis," *Proceedings of 8th Annual Reliability and Maintainability Conference*, Ame[illegible]stitute of Aeronautics and A[illegible]nautics, Denver, July 1969, 251-254.

174. Swain, A. D., "A Work Situation Approach to Improving Job Safety," *Proceedings, 1969 Professional Conference, American Society of Safety Engineers*, Chicago, August 1969, 233-257.

175. Swain, A. D., "Development of a Human Error Rate Data Bank," in J. P. Jenkins (ed), *Proceedings of U.S. Navy Human Reliability Workshop*, 22-23 July 1970, Naval Ship Systems Command, Office of Naval Research, and Naval Air Development Center, Washington, DC: February 1971, 113-148.

176. Swain, A. D., "Design of Industrial Jobs a Worker Can and Will Do," *Human Factors*, 1973, *15*, 129-136.

177. Swain, A. D., "Shortcuts in Human Reliability Analysis," in E. J. Henley and J. W. Lynn (eds), Generic Techniques in Systems Reliability Assessment, Leyden, The Netherlands: Noordhoff International Pub, 1974, 393-410.

178. Swain, A. D., Human Factors Associated with Prescribed Action Links, SAND74-0051, Sandia National Laboratories, Albuquerque, NM July 1974.

179. Swain, A. D., Preliminary Human Factors Analysis of Zion Nuclear Power Plant, NUREG76-6503, U.S. Nuclear Regulatory Commission, Washington, DC, October 1975.

180. Swain, A. D., Sandia Human Factors Program for Weapons Development, SAND76-0326, Sandia National Laboratories, Albuquerque, NM, June 1976.

181. Swain, A. D., "Error and Reliability in Human Engineering," in B. B. Wolman (ed), International Encyclopedia of Psychiatry, Psychology, Psychoanalysis, and Neurology, New York: von Nostrand Reinhold, Aesculapius Publishers, 1977, Vol. IV, 371-373.

182. Swain, A. D., "Estimating Human Error Rates and Their Effects on System Reliability," in Fiabilite et Disponibilite des Systemes Mecaniques et de Leurs Composants, Cycles de Conferences, Electricite de France - Commissariat a l'Energie Atomique, Jouy-en-Josas, France, October 1977.

183. Swain, A. D., The Human Element in Systems Safety, A. D. Swain, 712 Sundown Pl. SE, Albuquerque, NM 87108, May 1980 (Revised).

184. Swain, A. D., Design Techniques for Improving Human Performance in Production, A. D. Swain, 712 Sundown Pl. SE, Albuquerque, NM 87108, June 1980 (Revised).

185. Swain, A. D. and Bell, B. J., "Effects of Length of Work Period on Visual Detection and Decision-Making with Implications for Nuclear Power Plant Personnel," Nuclear Safety, 1980, in press.

186. Swain, A. D. and Guttmann, H. E., "Human Reliability Analysis Applied to Nuclear Power," Proceedings of 14th Annual Reliability and Maintainability Conference, Inst. of Electrical and Electronic Engineers, New York, January 1975, 116-119.

187. Swain, A. D. and Wohl, J. G., Factors Affecting Degree of Automation in Test and Checkout Equipment, DTA-TR-60-36F, Dunlap & Assoc., Darien, Conn., 1961.

188. Thurstone, L. L., The Measurements of Values, Chicago: Univ. of Chic. Press, 1959.

189 Torgerson, W. S., Theory and Methods of Scaling, New York: Wiley, 1958.

190. Trumbull, R., Analysis of Basic Factors Constituting Necessary Mathematical Proficiency Required for Success in Naval Aviation: Report I, Addition, Rpt. No. N.M. 001058.20.01, U.S. Naval School of Aviation Medicine, Naval Air Station, Pensacola, FL, 1952.

191. Trumbull, R., Analysis of Basic Factors Constituting Necessary Mathematical Proficiency Required for Succcess in Naval Aviation: Report II, Subtraction, Rpt. No. N.M. 001058.20.02, U.S. Naval School of Aviation Medicine, Naval Air Station, Pensacola, FL, 1952.

192. Trumbull, R., "Diurnal Cycles and Work-Rest Scheduling in Unusual Environments," Human Factors, 1966, 8, 385-389.

193. Van Cott, H. P. and Kinkade, R. G. (eds), Human Engineering Guide to Equipment Design (rev ed), Washington, DC: U.S. Government Printing Office, 1972.

194. Vernon, H. M. and Bedford, T., Two Studies of Absenteeism in Coal Mines I. The Absenteeism of Miners in Relation to Short Time and Other Variables, IHRD Report #62, Her Majesty's Stationery Office, London, 1931.

195. Vesely, W. E. and Goldberg, F. F., "Time Dependent Unavailability Analysis of Nuclear Safety Systems," IEEE Transactions on Reliability, 1977, R-26, 257-260.

196. WASH-1400 (NUREG-75/014): Reactor Safety Study -- An Assessment of Accident Risks in U.S. Commercial Nuclear Power Plants, U.S. Nuclear Regulatory Commission, Washington, DC, October 1975: Main Report, Appendix II - Fault Trees, Appendix III - Failure Data.

197. Webster's New Collegiate Dictionary, Springfield, MA: G. & C. Merriam Company, 1975.

198. Wechsler, D., Range of Human Capacities, Baltimore: Williams & Wilkins, Co., 1952.

199. Weldon, R. J., Error Reductions in Computations, SC-4074(TR), Sandia National Laboratories, Albuquerque, NM, Nov. 1956.

200. Weldon, R. J., Yafuso, R., and Peterson, G. M., Effect of Dial Design on Speed and Accuracy, II, Special Purpose Dials, SC-3839(TR), Sandia National Laboratories, Albuquerque, NM, 1955.

201. Welford, A. T., "Stress and Performance," in Welford, A. T. (ed), Men Under Stress, New York: Halsted Press, 1974.

202. Welford, A. T., Skilled Performance: Perceptual and Motor Skills, Glenview, IL: Scott, Forseman, 1976.

203. Wiener, N., Cybernetics, New York: Wiley, 1948.

204. Wolcott, J. H., McMeekin, R. R., Burgin, R. E. And Yanowitch, R. E., "Correlation of General Aviation Accidents with the Biorhythm Theory," Human Factors, 1977, 19, 283-293.

205. Woodson, W. E. and Conover, D. W., Human Engineering Guide for Equipment Designers, Berkeley, CA: Univ. of California Press, 1964, (2nd ed).

206. Woodworth, R. S., Experimental Psychology, New York: Henry Holt, 1938.

207. Zeff, C., "Comparison of Conventional and Digital Time Displays," Ergonomics, 1965, 8, 339-345.

GLOSSARY

Following are definitions of the technical terms used in the handbook and their abbreviations, if commonly used. A separate List of Abbreviations is also provided. The page number where each term first appears in the handbook is listed. Additional page numbers are listed where the term is defined or discussed more fully. In some cases it is convenient to use one of the defined terms from this glossary in defining another term. In such case, this referent term in underlined. The definitions are not intended to represent the last word in technical accuracy, but to provide the reader with enough understanding for purposes of this handbook. For more technical definitions of the psychological terms, see English and English (1958).

Abnormal Operating Conditions (2-26) - a general term to designate nonnormal plant conditions; e.g., the occurrence of a transient.

Accident-Prone Person (2-21) - one who statistically has "more than his share" of accidents compared with people with the same degree of exposure to accident-prone situations.

Accident-Prone Situation (APS) (2-21) - a situation that fosters human errors likely to result in injury to people or damage to equipment.

Administrative Control (15-1) - a general term referring to the kinds of checking of human performance mandated in a plant and the extent to which plant policies are carried out and monitored, including tagging control and lock and key control for valves and other components.

Annunciated Display (2-27, 10-1) - a legend indicator with an auditory alarm to announce that a change of state has occurred.

Annunciator Tile (2-27) - an individual annunciated legend indicator.

Anticipated Transient Events (10-22) - anticipated perturbations in the normal operating condition of a nuclear power plant that may require rapid reactor shutdown.

Anticipated Transient Without Scram (ATWS) (2-26) - an anticipated <u>transient</u> event not accompanied by an automatic <u>reactor trip</u> (or scram).

Approximate Failure Equation (7-40) - the sum of the end products of the failure paths in a system, assigning a probability of 1.0 to the success probabilities in these paths. For most reliability analyses, this approximation is sufficiently accurate when none of the failure probabilities of events in a path is greater than 10^{-2}.

Arousal - see <u>facilitative stress</u>.

Automatic Control (12-1) - an arrangement for the response of components or systems to signal from sensors or computers, without human intervention.

Auxiliary Feedwater System (AFWS) (11-3, 21-14) - a standby system to provide water to the secondary side of the steam generator in case the main feedwater system fails. Depending on the NPP, switchover to the AFWS may be automatic or manual.

Auxiliary Reactor Operator (3-14) - an unlicensed <u>reactor operator</u> trainee who assists the licensed operators.

Availability (2-2) - the probability that a system or component is available for use when needed.

Average (Industrial) Conditions (20-1) - conditions which do not subject a worker to an unusual degree of discomfort or distress.

Average Downtime (6-2) - the average time that transpires between the initial failure of a system or component and its return to an available state.

Balance of Plant Panels (3-26) - the panels of controls and displays for functions other than Reactor Control and Engineered Safety Features; e.g., power distribution.

Barrier Limits (2-16) - a type of <u>human tolerance limit</u> consisting of physical restraints that limit the range of a <u>control</u> or controlled element to prevent out-of-limits performance.

Basic Human Error Probability (BHEP) (2-11) - the probability of a human error on a task which is considered as an isolated entity; i.e., not influenced by previous tasks.

Basic Human Success Probability (BHSP) (7-7) - the complement of the basic human error probability.

Basic Walk-Around (8-1) - a type of relatively passive inspection, usually by an auxiliary reactor operator, in which he walks through the plant to detect any unusual condition, including deviant conditions of equipment. Any instructions to note a specific item of equipment are not part of the basic walk-around.

Basic Work Situation (for Valve Changing or Restoration) (13-3) - one in which a single valve is to be mainpulated, which is in a separate location from the work area, and is not adjacent to any similary valves.

Best-Case Analysis (II-3, 21-21) - an analysis in which consistently low estimates of human error probabilities (HEPs) (e.g., the lower uncertainty bound for each HEP) are used to present an overly optimistic assessment of the role of the human. Often this analysis is part of a bounding analysis.

Boiling Water Reactor (BWR) (2-22) - a type of light water reactor in which steam is generated in the reactor vessel and used to drive a turbine; thus, the water in the turbine is radioactive.

Bounding Analysis (II-3, 21-20) - an analysis in which best- and worst-case system reliability estimates are obtained by using, respectively, the lower and upper error bounds for the human components of the system. See best-case analysis and worst-case analysis.

Cause-Consequence Diagram (5-9) - a graphic representation of events in a system, combining elements of event trees and fault trees.

Caution Limit (2-17) - a human tolerance limit given by warnings, signs, or other indications.

Change (of valve status) (13-1) - change of the state of a valve from its normal operationing position to a nonnormal position.

Checker (13-9, 15-2) - one who is assigned to verify the accuracy of another's work, either while that person is doing the work or after its completion. The terms inspector and monitor are used interchangeably with checker.

Closed Loop System (3-1) - a system in which information about its outputs are fed back to become part of the system's inputs, so that system error can be responded to. Human error is part of the system error. This type of system can be contrasted with an "open loop" system in which this feedback is absent.

Common Cause Dependence (7-2) - A situation in which the performance of two or more tasks is related to some common influence; e.g., stress. Some writers employ the term "common mode."

Common-Cause Failure Event (5-14) - a failure in which common cause dependence results in the failure of two or more system components.

Commission Error - see Error of Commission.

Complete Dependence (CD) (between two tasks) (2-19, 7-17) - a situation in which, if the relationship between the tasks is positive, failure on one task will result in failure on the other with certainty. Similarly, if success occurs on one task, success will occur on the other. The opposite results will occur if the relationship between the tasks is negative.

Complete-Failure Path (5-7) - the only path through an event tree in which all tasks are performed incorrectly.

Complete-Failure Path Dependence (7-7) - the dependence among tasks that are all performed incorrectly.

Complete-Success Path (5-7) - the only path through an event tree in which all tasks are performed correctly.

Complete-Success Path Dependence (7-7) - the dependence among tasks that are all performed correctly.

Complexity Index (CI) (A14-10) - the average number of actions stated in the steps or paragraphs of written instructions.

Conditional Human Error Probability (CHEP) (2-11) - the probability of human error on a specific task given failure, or success, on some other task.

Consequences (to a system) (2-20) - the effects on a system of a human error that is not recovered.

Containment Sump (3-38) - bottom of the containment building in which the reactor vessel is located. Water from a loss-of-coolant accident flows to the sump by gravity.

Continuous Manual Controls (2-28, 12-1) - manual controls without detents which can be adjusted to any point within their range (e.g., a potentiometer).

Continuous Task (2-28) - a task which involves some sort of tracking activity such as monitoring a changing situation. The control action in a continuous task can be either continuous (as in rod control) or discrete (as in stopping a pump when the water reaches some level).

Controls - see manual controls.

Control Coding (12-3) - specific identification of a manual control (e.g., a switch) by the use of such cues as color, shape, size, or position.

Control Labeling (12-3) - labels which identify function, possible positions, and other information related to a manual control.

Conventional Limits (2-17) - a human tolerance limit in which the limits of variability in human performance are instilled by training or custom, often without further reinforcement in the work situation.

Coupling (2-19) - a term used in WASH-1400 in lieu of dependence.

Critical Incident Technique (CIT) (17-13) - a set of procedures for collecting direct observations of human behavior in such a way as to facilitate their usefulness in solving practical problems and developing broad psychological principles. To be critical, the incident must make a significant difference in the outcome of the behaviors. Thus, a critical incident may be negative or positive in terms of its influence.

Demand Probability (2-10) - the probability that a given human action will be performed correctly when required.

Dependence (between two tasks) (2-19) - the situation in which the probability of failure (or success) on one task is different depending on whether a success or failure occurred on the other task.

Dependent Tasks (2-11) - two tasks are dependent if the conditional human error probability of the second task is different, depending on whether the first task was performed successfully or incorrectly.

Derived Data (1-6) - Data on human performance which are extrapolated from related performance measures.

Derived Human Error Probability (HEP) (19-1) - estimated human error probabilities (HEPs) based on extrapolation from HEPs collected in different situations from the one of primary interest.

Design-Basis LOCA (3-57, 21-13) - a loss-of-coolant accident in which there is a guillotine break in a very large coolant pipe to the reactor vessel, with the plant at full power and the water level in the primary receptacle for emergency coolant at the lowest level allowed by the NRC technical specifications.

Deviant Condition (8-1) - an unacceptable condition of some component or function in an NPP.

Deviant Display (9-2) - a visual display showing an unacceptable indication.

Deviant Item (8-2) - any NPP component in an unacceptable condition.

Deviant Manual Control (9-2) - a manual control in an unacceptable position.

Direct Dependence (7-2) - the situation in which the outcome of one task directly affects the outcome of a second task.

Directly Related Displays (9-3) - two or more displays that are closely related by function; e.g., reactor vessel temperature and pressure.

Discontinuous Manual Controls - see discrete manual controls.

Discontinuous Task (2-28) - one in which each task element is a discrete step (e.g., a calibration task). Synonym: discrete task.

Discrete Manual Controls (2-28, 12-1) - Manual controls that have a fixed number of positions, such as switches with detents.

Discrete Task - see discontinuous task.

Discrimination (4-14) - the process of detecting differences among signals to the sense organs.

Display (2-27) - any instrument or device that presents information to any sense organ (visual, auditory, or other).

Disruptive Stress (3-53) - the bodily or mental tension resulting from the response to a stressor that threatens a person, frightens, worries, or angers him, or increases his uncertainty, so that usually he performs at a decreased level of effectiveness or efficiency.

Doubling Rule (17-21) - when an operator is required to take some corrective action in moderately to extremely high stress conditions with very limited time available to take the corrective action, if the first action is ineffective his nominal human error probability for each succeeding corrective action doubles (up to the limit of 1.0).

Empircal (or Measurement) Limit (2-17) - a human tolerance limit checked by observation or measurement during or after performance.

Engineered Safety Feature (ESF) (1-2) - a system, device, or component designed to limit the adverse consequences to an NPP or surrounding environment in the event of abnormal operation or conditions; e.g., an emergency core cooling system to keep the nuclear reactor covered with water should there be a break in a coolant pipe.

Engineered Safety Feature (ESF) Panels (3-26) - the control room panels which house most of the controls and displays related to ESFs.

Engineering Psychologist (2-2) - a person working in the human factors area, generally a person with an advanced degree in experimental psychology.

Ergonomics (2-1) - the discipline concerned with designing machines, operations, and work environments so that they match human capacities and limitations.

Ergonomist (2-2) - a synonym for Engineering Psychologist; this term is most frequently used outside of the U.S.

Error - see human error.

Error Bounds - see uncertainty bounds.

Error Correction (5-3) - the detection and correction of incorrect task performance in time to avoid any undersirable consequences to the system.

Error-Likely Person (2-21) - a person who consistently makes significantly more errors than others performing the same task (or tasks) under the same conditions, or a person who temporarily makes "more than his share" of errors due to temporary conditions such as fatigue, emotional upset, etc.

Error-Likely Situation (2-21) - a work situation in which the performance shaping factors are not compatible with the capabilities, limitations, or needs of the person performing a task.

Error of Commission (2-8) - incorrect performance of a task (or action).

Error of Omission (2-8) - failure to perform a task (or action).

Error-Prone Person - see error-likely person.

Error-Prone Situation - see error-likely situation.

Event Tree (5-4) - a graphic form of task analysis in which actions are designated by limbs in the tree, and the sequence moves forward in time. The event tree is an inductive model whereas the fault tree is a deductive model.

Experienced Person (18-1) - one who has at least 6 months' experience on the job for which he is qualified and/or licensed.

External Performance Shaping Factor (PSF) (3-4) - a performance shaping factor which is outside the individual and defines the work situation for him.

Extraneous Act (2-8) - a special category of error of commission whereby a person introduces some task or step that should not have been performed.

Extremely High Stress Level (3-55, 17-11) - the level of perceived stress that for most people will be extremely disruptive to system-required behavior.

Facilitative Stress (3-55) - the bodily or mental tension resulting from the response to a stressor that alerts a person, prods him to action, thrills him, or makes him eager, but not too much, so that usually he performs at a optimum level of effectiveness or efficiency.

Fault Tree (5-9) - a graphic representation of system events staring with some deviant condition and working backwards in time. The fault tree is a deductive model whereas the event tree is an inductive model.

Feedback (3-23) - the knowledge of results that one receives about the status or adequacy of his outputs.

First-Order Human Failure (5-1) - a situation in which a single human error will cause system failure. This is a type of single channel failure mode.

Fixed Limit (2-17) - a human tolerance limit which is clearly and permanently established; e.g., a red line on a meter to indicate safe maximum RPM.

Forensic Limit (2-17) - a human tolerance limit, usually determined after the fact, arrived at after debate (often legal) for the usual purpose of assigning blame for a human error that resulted in undesirable consequences.

Funneling of Attention (11-11) - the concentration on one source of information (e.g., a particular display) to the exclusion of other sources of information.

Halving Rule (17-22) - complement of the doubling rule whereby under an optimum stress level, a person takes extra care once he has made a mistake, and his human error probability (HEP) on his next attempt is half his nominal HEP for the task. This rule is not used in our analyses.

High Dependence (HD) (2-20, 7-17) - a level of dependence that is approximately mideway between zero and complete dependence on the continuum of dependence. (See Equations 7-12 and 7-17 in Table 7-1, p 7-30.)

High Pressure Injection System (HPIS) (5-13) - an engineered safety feature in a light water reactor to inject water into the primary loop when it is under high pressure, but losing relatively small amount of coolant, as from a small loss-of-coolant accident.

Human-Caused Error (HCE) (2-5) - an error whose primary causal factors are related to some human characteristic rather than to characteristics of the work situation.

Human Engineering (2-1) - see human factors.

Human Error (1-1, 2-4) - any member of a set of human actions that exceeds some limit of acceptability; i.e., an out-of-tolerance action, where the limits of human performance are defined by the system. Synonym: error.

Human Error Probability (HEP) (1-1, 2-9) - the probability that an error will occur when a given task is performed. Synonyms: human failure probability and task failure probability.

Human Error Probabiilty (HEP) Per Hour (2-10) - the calculated probability of at least one error (for a given task) per hour, regardless of the absolute frequency of errors in any given time period. This is a measure often used in unavailability calculations.

Human Error Rate (HER) (2-9) - the number of errors of a given type divided by the number of opportunities for the error. This term is not used in the handbook, but can be used interchangeable with human error probability.

Human Factors (2-1) - a discipline concerned with designing machines, operations, and work environments so that they match human capacities and limitations. Among human factors practitioners, the term is considered the general term which includes human factors engineering, procedures, training, selection, and any technical work related to the human factor in man-machine systems. In this handbook, the term is used interchangeably with human engineering, human factors engineering, and ergonomics.

Human Factors Engineering (2-1) - See human factors. Among human factors practitioners, the term is often restricted to design of equipment. In this handbook, the term is used interchangeably with human engineering, human factors, and ergonomics.

Human Factors Specialist (2-2) - a person working in the human factors area. Synonyms: engineering psychologist and ergonomist.

Human Failure Probability - see human error probability.

Human Performance Models (1-7, III-1) - descriptions of estimated human performance in a variety of NPP tasks presented as mathematical statements, with uncertainty bounds when appropriate. The models involve considerable extrapolation from available data and experience, and should be regarded as hypotheses.

Human Redundancy (2-20) - the error recovery factor resulting from the use of a person to check another's work.

Human Reliability (2-2) - the probability of successful performance of the human activities necessary for either a reliable or an available system, specifically, the probability that a system-required human act, task, or job will be completed successfully within a required time period, as well as the probability that no extraneous human actions detrimental to system reliability or availability will be performed.

Human Reliability Analysis (2-3) - a method by which human reliability is estimated.

Human Reliability Model (2-4) - a schematic representation or abstraction of human events and related system events and their interactions in a man-machine system. When probability values are assigned to the elements in the model, mathematical estimates of the probabilities of achieving (or not achieving) certain combinations of events in the system may be obtained.

Human Success Probability (HSP) (2-10) - the complement of human error probability; i.e., 1 - HEP.

Human Tolerance Limit (2-16) - a limit placed on human responses to keep the variability of human behavior within system-acceptable tolerances.

Importance of Effects (5-4) - generally a qualitative judgment of the relative importance of undesirable effects to a system in terms of cost or other system criteria.

Incredulity Response (3-57) - inability to accept or interpret evidence that some strongly undesired event has occurred, usually an unexpected event for which there has been little, if any, rehearsal of coping responses.

Independence (between two tasks) (7-1) - the situation in which the probability of failure or success on one task is the same regardless of whether failure or success occurred on the other. Synonym: zero dependence.

Independent Tasks (2-11) - two tasks are independent if the conditional human error probability of the second task is the same regardless of whether success or failure occurred on the other task.

Individual Performance Shaping Factors (PSFs) - see internal PSFs.

Initial Audit (9-2) - the oncoming shift operator's scanning of the control boards to assess the operating status of the plant and to detect any deviant conditions. Synonym: initial survey.

Initial Survey - see initial audit.

Inspection (20-13) - all monitoring activities for detecting conditons that are out of limits or approaching a limit. Inspection includes general observation of plant status, reading displays, and all forms of verification such as the status of switches, valves, and indicators.

Inspector (15-2) - one who inspects. Synonyms: checker and monitor.

Inspector Flinching (2-15) - an inspector's concern with only one limit of a range of acceptability, often resulting in a systematic error.

Intentional Error (2-6) - an error that occurs when the operator intends to perform some act that is incorrect but that he believes to be correct or to represent a superior method of performance. This type of error is not malevolent behavior.

Internal Performance Shaping Factors (PSFs) (3-4) - the characteristics of a human which affect his performance in a job, including personality characteristics, bodily structure, level of skill, attitudes, and so on.

Job and Task Instruction (3-4) - peformance shaping factors connected with the instructional materials used in jobs, and the manner in which job operations are intended to be carried out.

Joint Human Error Probability (JHEP) (2-11) - the probability of human error on all tasks which must be performed correctly to achieve some end result.

Labeling of Controls - see control labeling.

Licensee (A14-1) - the public utility that is licensed by the NRC to operate a nuclear power plant.

Licensee Event Report (LER) (3-20) - an event in an NPP which the NRC requires each licensee to describe. LERs are intended to include identification and evaluation of any human errors related to the reportable events.

Light Water Reactors (LWRs) (1-3) a type of nuclear power plant in which conventional water (as distinguished from heavy water) is used to remove the heat generated in a reactor vessel.

Link Analysis (2-26) - a form of task analysis in which movements of operating personnel (or movements of any bodily part; e.g., eyes) are plotted over some period of time.

Locally-Operated Valve - see manual value.

Loss-of-Coolant Accident (LOCA) (2-27) - a loss of reactor vessel coolant resulting from some defect such as a pipe break or leaky valve.

Low Dependence (LD) (between two tasks) (2-20, 7-16) - a level of dependence that is greater than zero dependence, but not very far up the continuum of dependence. (See Equations 7-10 and 7-15 in Table 7-1, p 7-30.)

Maintainer - see maintenance personnel.

Maintenance (of Valves) (13-1) - any testing, calibration, or other work requiring valves to be changed from their normal positions.

Maintenance Personnel (13-1) - NPP personnel who maintain or repair components such as valves and electrical or mechanical devices.

Malevolent Behavior (1-6) - deliberate behavior calculated to produce a harmful effect, thus, not a human error.

Man (2-2) - this term is used in its generic sense; i.e., pertaining to humans of either sex.

Man-Machine Interface (2-2) - any point of interaction between people and components in a system; e.g., a display, a manual control, or any other item a person observes or manipulates.

Man-Machine System (1-1, 2-2) - a system in which people have a monitoring and/or control function.

Man-Machine Systems Analysis (MMSA) (II-1) - a general method used to identify and evaluate existing or potential human performance problems in man-machine systems. The method includes task analysis and either qualitative or quantitative human reliability analysis.

Manual Control (2-28, 12-1) - the component with which the human enters his imputs to a system; e.g., switches, connectors, tools, etc.

Manual Valve (13-1) - a valve which is manually operated; in this handbook the term is restricted to locally-operated manual valves.

Measurement Limit - see empirical limit.

Mediating Processes (4-15) - the internal responses of a person, such as thinking, deciding, and worrying.

Mimic Lines (12-4) - lines on a panel to show the flow of energy or to indicate the desired sequence of human actions.

Model (2-4) - a model of a system is an abstraction that represents symbolically the way in which the system functions operationally; generally, not all characteristics of a system will be included in a model.

Moderate Dependence (MD) (2-20, 7-16) - level of dependence between low and high dependence. (See Equations 7-11 and 7-16 in Table 7-1, p 7-30.)

Moderately High Stress Level (3-58, 17-7) - the level of perceived stress that will be moderately disruptive to system-required behavior for most people.

Monitor (15-2) - one who inspects. Synonyms: checker and inspector.

Motor Control Center (MCC) (13-2) - a room in an NPP in which circuit breakers and parallel switches for certain control room functions are located.

Motor-Operated Valve (MOV) (2-28, 13-1) - a valve which is closed or opened by the action of a motor, usually controlled by a switch in the main control room.

Motor Responses (4-14) - those outputs of the operator with which he performs an action (usually with hands or feet).

Negative Dependence (7-3) - the situation in which failure on a task reduces the probability of failure on another task, or success on a task reduces the probability of success on another task. The handbook does not address negative dependence, but is restricted to positive dependence.

No-Cost Error (2-7) - a human error that does not result in undesirable system consequences.

Nominal Human Error Probability (HEP) (2-12, 5-12) - the central tendency of error probabilities for a task, used for single-point estimates of a human error probability.

Normal Control Room Operations (2-24) - planned operator tasks in the control room which include startup, shutdown, power level control, and calibration and testing.

Normal Operating Conditions (2-26) - a general term to designate plant conditions within normal limits, as distinct from abnormal operating conditions.

Normal Power Generating Condition (3-42) - a general term indicating that the plant is supplying power to the electric power grid. Synonym: steady-state operating mode.

Novice Personnel (18-1, 20-32) - NPP personnel with less than 6 months' experience in the job following qualification or certification.

Omission Error - see error of Omission.

One Annunciator (10-8, 20-8) - a single annunciator display or a functional group of annunciator displays with complete dependence among them, which are the equivalent of a single annunciator display.

One Display (9-5, 20-18) - a single display or functional group of displays with complete dependence among them, which are the equivalent of a single display.

One Item of Equipment (7-14, 8-14) - an individual item such as a display, control, valve, etc or some functional group of items that are completely dependent and are the equivalent of a single item with regard to errors of omission.

Operational Sequence Diagram (4-8) - type of event tree which presents information-decision-action sequences in a man-machine system.

Operator - as used in the handbook, the reactor operator in an NPP.

Optimum Stress Level (3-57, 17-6) - the level of perceived stress that is conducive to opimum performance. Most of the estimated human error probabilities in the handbook are predicated on the assumption of an optimum stress level.

Oral Procedures (14-1) - short verbal instruction given by someone in authority. See special instruction item.

Organismic Factors - see internal performance shaping factors.

Perceptual Set (7-2) - what a person expects to see or happen.

Performance Shaping Factors (PSFs) (2-18, 3-1) - internal or external factors that affect performance in a job context.

Physiological Stressor (2-23, 3-62) - stressors arising from physiological conditions such as fatigue, discomfort, high temperature, etc.

Populational Expectance (2-23) - a type of populational stereotype that refers specifically to displays which do not involve directional movement.

Populational Sterotype (2-23) - the way in which members of a population expect things to behave especially with respect to directional movements.

Position Indicator (20-22) - scale that indicates the postion of a valve relative to a fully opened or fully closed position.

Positive Dependence (7-4) - the situation in which failure on the first task increases the probability of failure on the second task, and success on the first task increases the probability of success on the second task. This is the only type of dependence considered in the handbook.

Pressurized Water Reactor (PWR) (3-26) - a type of light water reactor in which heat in the primary cooling loop is used to produce steam in a secondary loop; only the primary loop is radioactive.

Probability Density Function (pdf) (16-6) - in rough terms, a mathematical expression that gives the probability that a variable X will take values between two number X and X + X for all values of X. The pdf is often shown as a plot of events in a histogram representing proportionate frequency (i.e., the ratio of the number of events of interest to the total number of events) instead of actual frequency.

Probability Tree Digram (5-4) - an event tree in which the limbs designate human and other events as well as different conditions or influences upon these events. The values assigned to all the tree limbs (except those in the first branching) are conditional probabilities. The first limbs may also be conditional probabilities if they represent a carryover from some other tree. In any branching in the tree, the sum of the limbs is 1.0.

Psychological Scaling (19-5, 22-7) - techniques whereby the opinions of subject-matter experts are collected and pooled to determine the appropriate weights to assign to factors that influence human performance in a well-defined situation.

Psychological Stressor (2-23, 3-53) - stressor arising from external or internal factors that cause mental tension, (e.g., task load, threats, sensory deprivation, etc.). Psychological stressors can results in disruptive stress or facilitative stress.

Random Error (2-13) - out-of-tolerance actions that follow no predictable pattern but occur when the variability of behavior results in performance that is beyond system-acceptable variability.

Range Ratio (5-12) - the ratio of the highest score to the lowest score assigned to an event. In this handbook, a typical range ratio is the upper uncertainty bound of a human error probability (HEP) divided by the lower uncertainty bound of that HEP.

Reactor Control Panels (3-26) - panels housing the displays and manual controls associated with the reactor and associated components.

Reactor Operator (RO) (3-14) - personnel who are licensed to operate a control room in an NPP.

Reactor Trip (2-26) - a scheduled or unscheduled event in which the reactor vessel is rendered subcritical by insertion of all control rods. Synonym: scram.

Recovered Error (2-8, 4-19) - an error that was detected and corrected in time so that no undesirable consequences to the system were possible.

Recovery Factors (2-20) - factors which prevent or limit the adverse consequences of a human error.

Reference Limit (2-17) - a human tolerance limit consisting of a standard of acceptable performance with which a person can compare his own performance.

Refueling Water Storage Tank (RWST) (3-38) - a tank in a pressurized water reactor which holds part of the emergency core cooling water, and is also used to supply water to keep the fuel rods covered during refueling.

Reliability (2-2) - the probability of successful performance of function.

Response Perseveration (3-59) - the tendency to make some incorrect response (or a very limited number of such responses) repeatedly.

Restore (valve status) (13-1) - returning the state of a valve to its normal operating position.

Rule-Based Behavior (III-2) - behavior associated with tasks for which the appropriate actions are specified in step-by-step procedures; e.g., calibration and preventive maintenance tasks.

Scram - see reactor trip.

Senior Reactor Operator (SRO) (3-14) - a licensed reactor operator who has had sufficient experience and who has passed the examinations for the senior designation.

Sensitivity Analysis (II-3, 4-20, 7-41, 21-22) - an analysis in which one or more estimates of various parameters are varied to observe their effects on a system or some part of it (e.g., in a reliability analysis, estimates of human error probabilities would be varied to observe their effects on task or system reliability).

Sequential Error (2-8) - performance of some task or step out of the specified sequence.

Simple Multiplicative Model (5-8) - a performance model in which task probabilities are multiplied with the assumption that no dependence exists among the tasks. This model is rarely used in human reliability analysis.

Single Channel Failure Modes (5-14) - a situation in which one failure (e.g., a single human error) can result in system failure.

Single Point Human Error Probability (HEP) (5-11) - the use of a single point, ordinarily the nominal human error probability, to represent the entire distribution of human error probabilities for a task.

Situation-Caused Error (SCE) (2-5) - an error whose primary causal factors are related to the design of the work situation.

Situational Characteristics (3-4) - those performance shaping factors that are often plant-wide in influence, or that cover many different jobs and tasks in the plant; e.g., ambient temperature, peer relationships, etc.

Skill of the Craft (14-2) - a term describing those tasks in which it is assumed that the workers know certain aspects of the job and need no written instructions; e.g., a plumber replacing a washer in a faucet.

Special Instruction Item (14-1) - a specific or general item of instruction given in writing or orally; e.g., "restore the valves for System ABC."

Sporadic Error (2-15) - infrequent actions that are outside the tolerance limits for a system, despite small variability in performance.

Steady-State Condition (of a display) (9-4) - the situation in which displays are not rapidly changing, lamps are not blinking, and the auditory signals for annunciators are quiet.

Steady-State Condition (statistical) (6-2) - the probability of being in a failed state is independent of time.

Steady-State Operating Mode - see normal power operating condition.

Stress (2-23) - the human response to a stressor.

Stressor (2-23) - any external or internal forces that cause bodily or mental tension (i.e., stress).

Subject-Matter Expert (4-12) - experienced people on a job.

Systematic Error (2-14) - out-of-tolerance actions characterized by a performance dispersion pattern offset from a desired norm, indicating a consistent bias.

Tagging System (15-10) - all those administrative controls that ensure (1) awareness of any valves or other items of equipment that are in a nonnormal state, and (2) prompt restoration of this equipment to the normal state after the completion of test or maintenance operation. A tagging system includes the use of tags; chains, locks, and keys; and logs, suspense forms, and any other techniques that provide an inventory of the above items.

Talk-Through (3-68, 4-12) - a task analysis method in which an operator describes the actions required in a task, explains what he is doing and his mental processes during each action in actual or simlulated performance of a task. If the performance is simulated, the operator touches the manual controls he would operate and describes the control action required. He points to displays and states what readings he would expect. He describes any time delays and feedback signals, and the implications to the plant function of his actions. Synonym: walk-through.

Task Analysis (2-25, 4-4) - an analytical process for determining the specific behaviors required of the human components in a man-machine system. It involves determining the detailed performance required of people and equipment, and the effects of environmental conditions, malfunctions, and other unexpected events on both. Within each task to be performed by people, behavioral steps are analyzed in terms of (1) the sensory signals and related perceptions, (2) the decisions, memory storage, and other mental processes, and (3) the required responses.

Task and Equipment Characteristics (3-4) - those performance shaping factors that are specific to a given task and/or to the equipment required for completion of that task; e.g., the design of a tool, a display, or some other man-machine interface.

Task Effects (5-3) - generally the task effects of interest in a human reliability analysis are those undesirable consequences to a system resulting from incorrect and uncorrected task performance. Task effects can, of course, be positive.

Task Failure Probability (2-10) - used interchangeably with human error probability.

Task Reliability (5-3) - used interchangeably with human success probability.

Task Success Probability (2-10) - used interchangeably with human success probability.

Task Taxonomy (2-24) - a classification scheme for tasks.

Technique for Human Error Rate Prediction (THERP) (5-2) a method developed at Sandia National Laboratories to assess quantitatively the influence of human errors in a system. It is the method employed in this handbook.

Tiles - see annunicator tiles.

Time Error (2-8) - failure to perform a task or step within system-alloted time; i.e., completion of the actions either too early or too late.

Total Range Ratio (16-2) as defined by Wechsler (1952), the ratio of the highest score to the lowest score of a group of people homogeneous with respect to what is being measured, but excluding the extreme scores defined as the lowest and highest tenths of 1% of the population.

Transient (2-26) - an abnormal operating condition in which some NPP function departs from normal limits; e.g., loss of main feedwater.

Typical NPP Procedures (3-48, 13-3, 14-5) - written procedures that are narrative in style, with a low signal-to-noise ratio, and which require at least a Grade 12 reading level.

Unannunciated Displays (2-27) - meters, digital readouts, chart recorders, graphs, indicator lights, computer printouts, and video presentations not accompanied by auditory alerting signals.

Unanticipated Transient (10-22) - a transient which has not been anticipated in design of safety features in an NPP.

Unavailability (2-12, 6-1) - the probability that a system is not available for use when needed.

Uncertainty Bounds (2-11) - the upper and lower bound of human error probabilities (HEPs) that reflect the uncertainty in the estimation of a nominal HEP. The bounds include the variability of people and conditions and the uncertainty of the analyst in assigning HEPs to a task are judged to include the middle 90% of the HEPs for that task.

Unintentional Error (2-7) - a mistake that was not intended; e.g., dropping a tool, inadvertently tripping a switch, forgetting a step in a written procedure.

Unrecovered Error (2-8, 4-19) - an error which is not detected and corrected, and which could result in some undesirable consequence to the system.

Very Low Stress Level (3-55, 17-4) - the level of perceived stress that does not produce sufficient arousal to keep alert.

Vigilance Effect (3-60, 17-4) - the loss of alertness that results when a person is not sufficiently aroused. (Note: The handbook uses this term merely in the above descriptive sense.)

Walk-Around Inspection (8-1) - a scheduled inspection tour of a specified area in an NPP for which the inspector is required to report anything unusual or any deviant condition. In addition, he may be given some specific instructions; e.g., "Make sure that the main isolation valve on the RWST is open." (See also basic walk-around.)

Walk-Through - see talk-through.

Wechsler's Range Ratio (16-5) - Given a population that is homogeneous with respect to some ability that can be measured with a ratio scale, if the extreme scores are discarded (i.e., the lowest and highest tenths of 1%), the highest score in the distribution will rarely be more than five times the smallest score, and usually not more than three times the smallest score. This is similar to Wechsler's (1952) definition of total range ratio, except that this ratio can also apply to distributions of human error probabilities, as discussed in the text.

Work Situation Approach (WSA) (2-6) - an approach to identifying and analyzing error-likely situations in which it is assumed that the primary causal factors behind most human errors in a well-structured work situation are more closely related to such system elements as operating procedures, equipment design, and management practices than to the individual characteristics of trained personnel.

Worst-Case Analysis (II-2, 21-21) - an analysis in which consistently high estimate of human error probabilities (HEPs) (e.g., the upper uncertainty bound for each HEP) are used, yielding an overly pessimistic assessment of the role of human error. Often this analysis is part of a bounding analysis.

Zero Dependence (ZD) (between two tasks) (2-19, 7-14) - the situation in which the probability of failure or success on one task is unaffected by failure or success on the other. Synonym: independence.

ABBREVIATIONS

Following is a listing of the abbreviations used in the handbook. For definitions of any of the terms, see the Glossary.

AC	Aircraft Commander
AFWS	Auxiliary Feedwater System
ANN	Annunciator
APS	Accident-Prone Situation
ASME	American Society of Mechanical Engineers
ATWS	Anticipated Transient Without Scram
BHEP	Basic Human Error Probability
BHSP	Basic Human Success Probability
BWR	Boiling Water Reactor
CD	Complete Dependence
CHEP	Conditional Human Error Probability
CHRS	Containment Heat Removal System
CI	Complexity Index
CIT	Critical Incident Technique
CLS	Consequence Limiting System
DO	Dedicated Operator
ECCS	Emergency Core Cooling System
ELS	Error-Likely Situation
EPRI	Electric Power Research Institute
ESF	Engineered Safety Feature
FSAR	Final Safety Analysis Report
HCE	Human-Caused Error
HD	High Dependence
HEP	Human Error Probability

HER	Human Error Rate
HPCI	High Pressure Containment Injection
HPIS	High Pressure Injection System
IEEE	Institute of Electrical and Electronic Engineers
IREP	Interim Reliability Evaluation Program
JHEP	Joint Human Error Probability
LB	Lower Bound
LD	Low Dependence
LER	Licensee Event Report
LOCA	Loss-of-Coolant Accident
LWR	Light Water Reactor
MCC	Motor Control Center
MD	Moderate Dependence
MMSA	Man-Machine Systems Analysis
NPP	Nuclear Power Plant
NRC	Nuclear Regulatory Commission
OSHA	Occupational Safety and Health Act
pdf	Probability Density Function
PSF	Performance Shaping Factor
PWR	Pressurized Water Reactor
QC	Quality Control
RO	Reactor Operator
RWST	Refueling Water Storage Tank
SCE	Situation-Caused Error
SD	Standard Deviation
SI	Safety Injection
THERP	Technique for Human Error Rate Prediction
TMI	Three Mile Island

UB	Upper Bound
U.S.	United States
WSA	Work Situation Approach
ZD	Zero Dependence

EQUATIONS

Following are the equations used in the handbook. The equations are listed by equation number and page number.

Equation Number	Equation	Page
(6-1)	$A = \frac{\bar{u}}{\bar{u}+\bar{d}}$	6-1
(6-2)	$U = 1 - A = 1 - \frac{\bar{u}}{\bar{u}+\bar{d}} = \frac{\bar{d}}{\bar{u}+\bar{d}}$	6-1
(6-3)	$U = \frac{p\bar{d}}{\bar{T}}$	6-2
(6-4)	$p = ER$	6-2
(6-5)	$\bar{d}_T = h_1 + C_1h_2 + C_1C_2h_3 + \ldots + C_1C_2\ldots C_m\bar{T}$	6-3
(7-1)	$Pr[S\|ZD] = ab\ldots n$	7-25
(7-2)	$Pr[F\|ZD] = AB\ldots N$	[illegible]
(7-3)	$Pr[S\|ZD] = a^{\underline{n}}$	7-25
(7-4)	$Pr[F\|ZD] = A^{\underline{n}}$	7-25
(7-5)	$Pr[S\|CD] = a$	7-25

Equation Number	Equation	Page
(7-6)	$Pr[F \mid CD] = A$	7-25
(7-7)	$Pr[S \mid CD] = \frac{a + b + \ldots + n}{\underline{n}}$	7-26
(7-8)	$Pr[F \mid CD] = \frac{A + B + \ldots + N}{\underline{n}}$	7-26
(7-9)	$Pr[S_{\text{"}N\text{"}} \mid S_{\text{"}N-1\text{"}} \mid ZD] = n$	7-30
(7-10)	$Pr[S_{\text{"}N\text{"}} \mid S_{\text{"}N-1\text{"}} \mid LD] = \frac{1 + 19n}{20}$	7-30
(7-11)	$Pr[S_{\text{"}N\text{"}} \mid S_{\text{"}N-1\text{"}} \mid MD] = \frac{1 + 6n}{7}$	7-30
(7-12)	$Pr[S_{\text{"}N\text{"}} \mid S_{\text{"}N-1\text{"}} \mid HD] = \frac{1 + n}{2}$	7-30
(7-13)	$Pr[S_{\text{"}N\text{"}} \mid S_{\text{"}N-1\text{"}} \mid CD] = 1.0$	7-30
(7-14)	$Pr[F_{\text{"}N\text{"}} \mid F_{\text{"}N-1\text{"}} \mid ZD] = N$	7-30
(7-15)	$Pr[F_{\text{"}N\text{"}} \mid F_{\text{"}N-1\text{"}} \mid LD] = \frac{1 + 19N}{20}$	7-30
(7-16)	$Pr[F_{\text{"}N\text{"}} \mid F_{\text{"}N-1\text{"}} \mid MD] = \frac{1 + 6N}{7}$	7-30

Equation Number	Equation	Page
(7-17)	$\Pr[F_{\text{"N"}} \mid F_{\text{"N-1"}} \mid HD] = \frac{1+N}{2}$	7-30
(7-18)	$\Pr[F_{\text{"N"}} \mid F_{\text{"N-1"}} \mid CD] = 1.0$	7-30
(8-1)	$\Pr[S_{\leq 30\text{ days}} \mid \text{1 shift, 1 inspector}]$ $= 1 - \Pr[F_{\text{Day 1}}]\Pr[F_{\text{Day 2}}]\Pr[F_{\text{Day 3}}]\Pr[F_{\text{Day 4}}]\Pr[F_{\text{Day 5}}]^{26}$ $= 1 - (.9 \times .95 \times .975 \times .99 \times .999^{26})$ $= .1959 \approx .20$	8-6
(8-2)	$\Pr[S_{\leq 30\text{ days}} \mid \text{3 shifts, 3 inspectors} \mid \text{ZD between shifts}]$ $= 1 - \{1 - \Pr[S_{\leq 30\text{ days}} \mid \text{1 shift, 1 inspector}]\}^3$ $= 1 - .8041^3 = .480 \approx .5$	8-6
(8-3)	$\Pr[S_{>30\text{ days}} \mid F_{30\text{ days}}] = 0$	8-7
(8-4)	$\Pr[S_{\leq 30\text{ days}} \mid \text{1 shift, 30 inspectors}]$ $= 1 - \Pr[F_{\text{Day 1}}]^{30}$ $= 1 - .9^{30} \approx .96$	8-8
(8-5)	$\Pr[S_{\leq 30\text{ days}} \mid \text{3 shifts, 90 inspectors}]$ $= 1 - \{1 - \Pr[S_{\leq 30\text{ days}} \mid \text{1 shift, 30 inspectors}]\}^3$ $= 1 - .042^3 \approx .9999$	8-8
(8-6)	$\Pr[S_{\leq 30\text{ days}} \mid \text{1 shift, 2 inspectors, each making 1 inspection per shift}]$ $= 1 - \Pr[F_{\text{Inspector 1}} \mid \text{30 days}]\Pr[F_{\text{Inspector 2}} \mid \text{30 days}]$ $= 1 - .8041^2 \approx .35$	8-13

Equation Number	Equation	Page
(8-7)	$\Pr[S_{\leq 30\text{ days}} \mid \text{3 shifts, 6 inspectors, 2 per shift, each making 1 inspection per shift}]$ $= 1 - \Pr[F_{\leq 30\text{ days}} \mid \text{1 shift, 2 inspectors, each making 1 inspection per shift}]^3$ $= 1 - .6466^3 \approx .73$	8-13
(8-8)	$\Pr[S_{\text{for any deviant item}} \mid \text{1/2 proper use and 1/2 improper use of checklist}]$ $= .5 \times (.99 + .9) = .945 \approx .95$	8-20
(8-9)	$\Pr[S_{\text{all } \underline{n} \text{ deviant items}} \mid \text{1/2 proper use and 1/2 improper use of checklist}] = .5 \times (.99^{\underline{n}} + .9^{\underline{n}})$	8-20
(8-10)	$D = .5 \times (1 - \Pr[F_{\underline{n}}])$	A8-1
(8-11)	$\Pr[F_{\underline{n}+1}] = \Pr[F_{\underline{n}}] + D \times (1 - r)$	A8-2
(9-1)	$\Pr[S_{\text{one or more deviant displays}} \mid \text{equal BHEPs}]$ $= 1 - \Pr[F_{\text{one}}]\left[\dfrac{1 + \Pr[F_{\text{one}}]}{2}\right]^{\underline{n}-1}, \ \underline{n} \leq 5$	9-13
(9-2)	$\Pr[F_{\text{both operators}}] = .5 \times \Pr[F_{\text{both operators}} \mid \text{MD}]$ $+ .5 \times \Pr[F_{\text{one operator}}], \ \underline{n} \leq 5$	9-16
(10-1)	$\Pr[F_i] = \begin{cases} 10^{-4}, \ i = 1 \\ 2^{i-2} \times 10^{-3}, \ 1 < i \leq 10 \\ .25, \ i > 10 \end{cases}$	10-18

Equation Number	Equation	Page
(10-2)	$\overline{\Pr[F_i]} = \sum_{1}^{n} \frac{\Pr[F_i]}{n}$	10-18