NUREG/CR-3837
BNL-NUREG-51786

MULTIPLE-SEQUENTIAL FAILURE MODEL: EVALUATION OF AND PROCEDURES FOR HUMAN ERROR DEPENDENCY

P.K. Samanta, J.N. O'Brien, and H.W. Morrison

Date Published — May 1985

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

Prepared for the U.S. Nuclear Regulatory Commission
Office of Nuclear Regulatory Research
Contract No. DE-AC02-76CH00016

NUREG/CR-3837
BNL-NUREG-51786
RX,AN,RG,1S

MULTIPLE-SEQUENTIAL FAILURE MODEL: EVALUATION OF AND PROCEDURES FOR HUMAN ERROR DEPENDENCY

P.K. Samanta, J.N. O'Brien, and H.W. Morrison*

Manuscript Completed -- March 1985
Date Published — May 1985

*Department of Psychology, State University of New York at Stony Brook, Stony Brook, NY 11794

Prepared by
ENGINEERING ANALYSIS AND HUMAN FACTORS GROUP
DEPARTMENT OF NUCLEAR ENERGY, BROOKHAVEN NATIONAL LABORATORY
UPTON, LONG ISLAND, NEW YORK 11973

Prepared for
UNITED STATES NUCLEAR REGULATORY COMMISSION
OFFICE OF NUCLEAR REGULATORY RESEARCH
WASHINGTON, D.C. 20555
UNDER INTERAGENCY AGREEMENT DE-AC02 76CH00016
FIN NO. A-3219

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

ABSTRACT

This report provides an evaluation of the practicality, acceptability, and usefulness of using the Multiple Sequential Failure (MSF) model originally described in NUREG/CR-2211, 1981. The MSF model is described, discussed, and procedures for its use provided. The model was found to be practical, acceptable, and useful as a PRA tool for assessing the degree of dependence due to human interactions with components in systems employing redundant components.

CONTENTS

CONTENTS (CONTINUED)

ACKNOWLEDGMENTS

The authors would like to acknowledge Dr. Thomas G. Ryan of the U.S. Nuclear Regulatory Commission who provided close guidance and substantial comments which shaped this report. Mr. William J. Luckas, Jr., Project Manager at BNL, provided support and valuable comments throughout the project. Mr. A. Azarm and Dr. P. Fahrazad provided significant help in developing the MSF computer code presented in this report. In addition, all of the individuals who reviewed the MSF model and this report are gratefully acknowledged for their assistance. Finally, the authors acknowledge Sharon M. Moore who provided invaluable clerical assistance and patience in typing the various drafts which preceded this report.

EXECUTIVE SUMMARY

This report provides a description of and users' guide for the Multiple Sequential Failure (MSF) model. The MSF model was initially developed and described in NUREG/CR-2211, (1981) for use as a tool in Probabilistic Risk Assessments (PRA) of nuclear power plants. The research effort described in this report was conducted in order to evaluate its practicality, usefulness, and acceptability to the PRA community as a methodological tool, as well as to make it generally available for use by PRA practitioners.

The MSF model can be used in a PRA to estimate the level of dependence in the failure of redundant components in a system due to human interactions. That is, because nuclear power plants employ redundant (i.e., backup) components in systems critical to plant safety as a means of decreasing the probability of total system failure, it is important to estimate the level of dependence (i.e., lack of independence) induced by maintenance, testing, and calibration activities which can cause common mode failures among redundant components. These service activities are often carried out in sequence, by the same personnel, using the same procedures and similar work practices. A common mistake made on each redundant component can usually cause total system failure. The model assumes that if a mistake is made in maintaining, testing, or calibrating one component, the same mistake is more likely to be made on the next redundant component to be serviced than if no previous mistake were made. For instance, if following testing of one pump in a system, the operator fails to restore a test-related valve to the correct position, the likelihood that he will fail to restore the same valve on the second pump as well increases if these tasks are performed sequentially.

The MSF model can be used to estimate the probability of multiple human failure in a system. The model uses the independent failure probability for a single component and systematically modifies it (i.e., increases it) for each redundant component in a sequence. The first component is assumed to have the same failure probability as that estimated for the component to fail randomly-due to human error. The second component is assumed to have a higher failure

probability than random if there was a human error committed during the servicing of the first component. If there are three or more components, the probability of failure is assumed to increase as human errors occur on the other components previously serviced, so that the probability of component failure increases sequentially. By inputting either actuarial or estimated data on failures of redundant components to the MSF computer code, the MSF model systematically modifies the random failure probabilities for each redundant component and then applies the failure logic of the system to arrive at a total system failure probability which accounts for dependence.

This report contains a discussion of the background on the development of the MSF model, a users' guide that can be employed for estimating human error dependency in conducting a PRA, the computer code for batch application of actuarial or estimated failure data, and a description of a small-scale psychological experiment. This experiment was used to determine the degree to which the MSF model accurately assesses dependent sequential errors committed by human subjects in analogous tasks. The experimental data collected were in general agreement with the MSF model.

1.0 INTRODUCTION

1.1 Purpose of This Report

The purpose of this report is to assess the practicality, usefulness and acceptability of employing the Multiple Sequential Failure (MSF) model (NUREG/CR-2211, 1981) as a tool in Probabilistic Risk Assessments (PRAs) of nuclear power plants. The MSF model is a tool for quantifying the human reliability component of system failure probabilities in safety-related systems which use redundant (i.e., backup) components. Because redundancy is employed in many of the most important safety-related systems, the calculation of failure probabilities due to human interactions in testing, maintenance, and calibration, must account for the effects of redundancy. These effects must be accounted for to ensure the accuracy and reliability of an overall risk assessment.

1.2 Overview

The conservative safety philosophy employed in the design of nuclear power plants requires the use of redundancy to assure that the probability of reactor accidents remains acceptably low. If failures of redundant components in safety-related systems are assumed to be independent, the probability of total system failures will be low. However, actual experience in the operation of power plants has revealed that certain total system failures events occur more frequently than would be pre cted using the assumption of independence. This can be explained by considering sources of dependence in failures among redundant components as a result of physical and human interactions. Dependence can be defined as the condition where the probability of a redundant component failure is modified (usually increased) away from its independent failure probability because of another previous failure (i.e., the probability of each successive component failure in a redundant system is increased). Epler (1969) and Fleming et al. (1983) provide numerous examples of dependence in system failures taken from actual reactor operating experience.

As discussed previously, analysis of dependent failures plays a major role in the estimation of risk from reactor accidents using PRA techniques. As a result, both qualitative and quantitative analyses of how dependent failures occur have received wide attention among practitioners conducting PRAs of nuclear power plants. The PRA Procedures Guide (NUREG/CR-2300, 1982) documents the various types of dependency analyses used in PRAs. It also recognizes that different quantification methods are needed for different types of dependent failures.

Human interactions with systems using redundant components are a dominant source of dependence in multiple failures of redundant components. In particular, testing, maintenance, and calibration activities provide a common link to redundant components in a system through human interactions. As a result, failures of separate redundant components in a system cannot be considered completely independent. Determination of the extent to which the probability of human error in one task is dependent on an error in a previous task is important in the realistic assessment of joint failure probabilities for redundant components because human interactions such as maintenance, testing, or calibration are usually performed sequentially. The MSF model was developed (1) to account for the type of human interactions that occur in nuclear power plants and (2) to be useful in the overall PRA process. This report details the MSF model and provides a procedure for calculating the joint human error probability for two or more redundant components while performing testing, maintenance, and calibration activities.

1.3 Relationship to General Human Reliability Methods

Human Reliability Analysis (HRA) is an important element in PRAs. HRA considers and quantifies probabilities of human error in tasks that are performed both under normal conditions and during accidents or abnormal occurrences. Under normal conditions, the errors which are considered usually occur during or after testing, maintenance, or calibration; under accidents or abnormal conditions, the errors considered are associated primarily with operator responses to the indication of conditions in the control room. A major problem in HRA analysis, as

discussed above, is the determination of dependency among these errors. The MSF model is a method specifically developed for calculating dependency during normal conditions (i.e., from maintenance, testing, and calibration).

NRC research and development in the area of human reliability in nuclear power plants has in part led to the development of the Handbook of Human Reliability Analysis with Emphasis on Nuclear Power Plant Applications (NUREG/CR-1278, 1983). This Handbook employs a general method called the Technique for Human Error Rate Prediction (THERP). The specific method recommended in the Handbook for using THERP in determining the effect of dependence among human errors is based primarily on the informed subjective assessment of the individual HRA specialist performing the analysis. THERP has been used effectively in several PRAs sponsored by NRC and utilities. Other general HRA methods available include Operator Action Trees (OATs) (NUREG/CR-3010, 1982) approach, used in the Susquehana Nuclear Plant Accident Initiation and Progress Analysis (AIPA) - and the Operator Response Model (Fleming et al., 1978), used in the High Temperature Gas Cooled Reactor (HTGR) risk assessment. Other HRA tools are generally discussed in NUREG/CR-2815, (1985).

The MSF model, which provides a unique method for quantification of dependent failures in HRA used in performing PRAs of nuclear power plants, can be used in conjunction with any of the general methods discussed above. This model addresses only a specific aspect of HRA analysis and is not a general method. It is most appropriate for determining dependent failure probabilities resulting from testing, maintenance, and calibration activities on systems that incorporate redundant components.

1.4 Organization of the Report

Section 2 briefly describes the basic features of dependent failures and the MSF model. It also discusses the advantages and disadvantages of using this method. Section 3 describes how the model is used in a risk assessment. Section 4 evaluates the major issues in evaluating the use of the MSF model, including those issues relating to the practicality, usefulness, and acceptability of

employing the model. Conclusions and recommendations are provided in Section 5. Appendix A provides a users' manual and describes the MSF model in detail, Appendix B describes a small-scale psychological experiment aimed at testing the model, and Appendix C lists the computer program for performing the calculations required to use the model.

2.0 THE CONCEPT OF DEPENDENT MULTIPLE SEQUENTIAL FAILURES AND THE MSF MODEL

Since redundancy is incorporated in safety-related systems in a nuclear power plant, the introduction of dependence among failures of redundant components in such a system will have substantial effects on the overall reliability of the plant. As a result of the importance to safety of systems which typically incorporate redundancy, the degree to which the probabilities of failure for redundant components are actually independent of each other is very important in determining overall risk from plant operation.

Examples of how human interactions may affect the independence of failure probabilities among redundant components in a system (i.e., introduce dependence) are (1) the same work crews performing an incorrect maintenance task on all redundant components in a system, (2) the same incorrect calibration of component controls due to an incorrectly set calibration instrument being used for work on redundant components in a system, (3) the use of incorrect procedures for testing all redundant components in a system, and (4) failure to restore critical valves on all redundant components after testing or maintenance of a system. Any of these examples will cause the overall system failure probability to be higher than if all redundant component failures are treated as being completely independent of each other.

2.1 Multiple Sequential Failure (MSF) Model

The MSF model developed by Samanta and Mitra (NUREG/CR-2211, 1981) provides a specific mathematical model and a statistical framework for quantifying the failure probability of systems employing redundancy due to human interactions associated with testing, maintenance, and calibration activities. A complete discussion of how the MSF model compares to other dependency modeling techniques used in risk assessments is provided in Appendix B.

As an example of how to use the model, it can be used to calculate a system failure probability considering human interactions with a system employing three redundant pumps. Ideally, human error data on this type of

system may be available. These redundant failure data, if complete, would include the number of times the set of pumps have been maintained, the number of times without any errors, the number of times only one error (and, therefore, pump failure) occurred, the number of times two sequential errors occurred, and the number of times three sequential errors (i.e., total system failure) occurred. A sample of these data is given in Table 2.1.

Table 2.1 Sample Human Error Data.

Number of Times:	Column 1 Failure Data Assuming Complete Dependence	Column 2 Failure Data Assuming Complete Independence
All Three Pumps were Maintained	1000	1000
No Pumps Failed due to Maintenance Error	980	987
One Pump Failed due to Maintenance Error	10	10
Two Pumps Failed due to Maintenance Error	7	3
Three Pumps Failed due to Maintenance Error	3	<1

The data in the first column of Table 2.1 can be used to calculate the level of dependence among these tasks. First, the single failure probability is approximately one in one hundred ($p \approx 0.01$). This yields a total of 10 single pump failures in 1000 maintenance tasks. If the pump failures were truly independent the probability of two pumps failing, which can happen in three different ways would be ($p = 0.0003$). This means that in the case illustrated in Table 2.1, that the failure of two pumps would have occurred approximately three times as indicated in the second column. For the failure of all three pumps, the probability it would be one in one million ($p = 0.000001$) and, therefore, out of 1000 maintenance actions, failure of all three pumps should occur less than one time as indicated in the second column. It is clear that the failure data in column 1 reflects a level of dependence because more multiple failures have occurred than would have if all three redundant pumps were completely independent.

Figure 2.1 describes the failure states in a 3-component system due to human errors during testing, maintenance, and calibration. In a 3-component system, one may observe the failure of one component, two component, and all three components. Since personnel at power plants perform their tasks on each component in a sequence, the probability of a total system failure due to human error depends to a great degree on which component fails first. Using the assumption that human interactions take place sequentially, an error committed on component A increases the probability of the same error being committed on components B and C. However, an error on component B and not on component A influences only component C because of the sequential procedure followed. Once a task has been performed on component A correctly, there is no reason to go back to component A; and if so, the entire process begins again. Thus, a failure caused by an error on component B can result in only two component failures (i.e., components B and C), and an error in component C will cause only the failure of component C. Determination of the probability of these multiple sequential failures, and therefore, total system failure, requires that the degree of dependency in this situation be assessed.

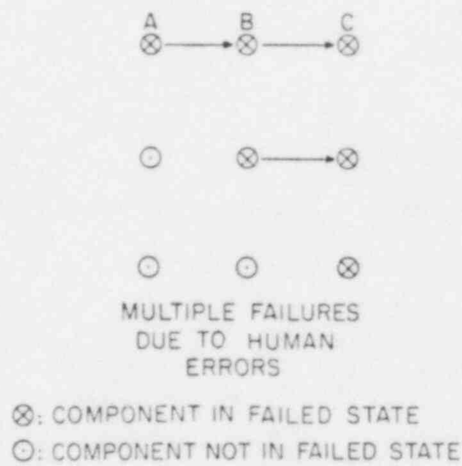


Figure 2.1 Effect of dependency in failure states resulting from human errors during testing, maintenance, and calibration activities in a 3-unit system.

Using the dependent failure data set in column 1 of Table 2.1, the independent failure probability of the first pump can be calculated. The second

pump's failure probability can be calculated assuming that the first failed and the third pump's failure probability can be calculated assuming failure of the first and second pumps. Using these derived failure probabilities, this model then arrives at an overall system failure probability which accounts for the dependence due to human interactions.

The MSF model uses two parameters to calculate probabilities of each component failing (i.e., the probabilities of sequential failures assuming one failure has already occurred). It uses the failure probability of the first pump failing (p) and a dependency factor (k). This can be illustrated by using the following equation for the conditional failure probability of the second pump for a system employing two redundant pumps:

$$p_{(\text{for pump 2 failing, assuming pump 1 has failed})} = p_{(\text{for pump 1})}{}^{+f(k,p)},$$

where f(k,p) indicates a function of k and p and is added to the probability of the first pump failing. The variable k can be calculated using human error data similar to those in column 1 of Table 2.1 by using the MSF computer program provided in Appendix C. These data can be simply input to the computer program and the resultant joint failure probability will be output.

In cases where those data are available, the overall failure probability of the multiple, redundant units in a system are calculated. If the only dependent failure data available are the number of errors on different combinations of components (i.e., the number of errors on one component, two components, and three components, in the above example), the probability of independent errors (i.e., of error in a single unit) can be used to estimate the number of opportunities for that failure to occur. This situation is realistic if, for example, the data are gathered from Licensee Events Reports (see NUREG/CR-3519, 1983 for an alternative method of calculating opportunities for errors). Limited data from available sources can be used in this way to generate data from which to calculate k and determine the multiple sequential failure probability for the overall system.

If no dependent failure data are available, then psychological scaling techniques such as structured expert judgment can be used to develop a data base upon which k can be calculated. NUREG/CRs-3688 and -4016 provide methods of using structured expert judgment which have been evaluated by NRC and are completely acceptable for the purpose of generating dependent failure data if they are not available.

The simplified discussion above can be generalized to present the four steps used by the MSF model in determining the multiple sequential failure probability for systems incorporating redundancy. First, the model defines the conditional failure probabilities expressed in terms of the individual conditional failure probabilities for each component in the system and the degree or level of dependence for each which may cause sequential failures. Second, on the basis of the number of components involved and the failure logic configurations of them, the model provides a mathematical expression for the multiple sequential failure probability (i.e., total system failure) due to human interactions during testing, maintenance, and calibration of multiple components. Third, the parameters of the model (independent failure probability [p] and the degree of dependence [k]) are either obtained from available dependent failure data or can be developed on the basis of structured expert judgment. Finally, multiple sequential failure probability is calculated using the expression developed in the second step and the estimated parameters in the third step.

2.2 General Features of the MSF Model

The details of the MSF model are discussed in NUREG/CR-2211 (1981). Also Appendix A of this report provides a users' manual and an overview of the major mathematical formulas used. In addition, Appendix B provides a complete discussion of how the MSF model compares to other dependency models. However, specific features of the model are briefly discussed below:

1. The MSF model provides the failure probability of multi-component system for various reliability configurations. For example, in a

four-component system, the model can be used to calculate the probability of failure of all four components, or of any three of the four components, or of any two, and so on, depending on how many components must fail for the overall system to fail.

2. The joint failure probability for multiple unit systems is calculated using two parameters: (1) the independent failure probability (p) and (2) the dependence factor (k) which is the degree of dependence induced by human interactions with the system. Thus, the model requires the calculation or estimation of these two parameters either from dependent failure data or through structured expert judgment.

3. The degree of dependence (k) is assumed to be a continuum, i.e., it can assume <u>any</u> value between complete independence (i.e., k = 0) to complete dependence (k = 1). The model calculates the degree of dependence (k) from a complete set of dependence failure data using statistical estimation procedure (moment estimation technique). However, the major problem in quantitative assessment of dependency among redundant components has been the lack of complete data. In situations where the actuarial data available on dependent failures are unavailable or considered insufficient, structured expert judgment may be used in two ways to determine multi-component failure probabilities (1) to estimate the degree of dependence among components directly or (2) to generate the number of times groups of components fail (i.e., a complete set of dependent failure data) for use in determining dependency using the MSF model and, thus, system failure probabilities.

The MSF model will calculate the conditional probabilities based on the estimated degree of dependence from structured expert judgment and then calculate the multi-component failure probabilities based on the system failure logic configurations. However, dependence is a complicated phenomenon and the degree of dependence may not be

sufficiently understood for an expert group to assess directly. Accordingly, the second, and perhaps preferable, way to introduce structured expert judgments is to estimate the number of times each of 1, 2, 3... n components may fail in an n-component system. For example, it may be estimated using structured expert judgment that in a 3-component system, out of 1000 maintenance interactions one of the components is expected to fail 12 times, two of the components five times, and all three of the components three times. The MSF model will then estimate the degree of dependence (k) from such a structured judgment and will, accordingly, calculate the failure probability of the multi-components system. Thus, in the case that insufficient actuarial data on dependence are available, the MSF model allows the incorporation of structured expert judgments and provides an effective means of using structured judgments in dependency evaluations.

4. An error in one task is assumed to increase the probability of an error in the next or subsequent task. This is positive dependence as opposed to negative dependence, in which case an error in one task decreases the probability of an error in the next task. The MSF model considers only positive dependence based on the need for conservatism as a specific assumption used in a PRA. Details on how positive dependence is modeled are provided in Appendix A and in NUREG/CR-2211 (1981).

5. The computations using the moment estimation technique to estimate k from a complete set of dependent failure data may become complicated and accordingly, a MSF computer program is provided (Appendix C) for the user to evaluate the failure probability of the multi-unit system directly.

6. The MSF model is used to analyze the specific aspect of Human Reliability Analysis (HRA) related to the quantification of dependence in systems employing redundant components due to human interactions.

The model can be used in conjunction with any appropriate general HRA method (e.g., THERP, OATS, etc.) in performing a PRA. The HRA analyst will be faced with situations in fault trees requiring the evaluation of multi-component failures due to human interactions in testing, maintenance, and calibrations. It is expected that the MSF model will be used as a tool to quantify system failure probabilities in those situations within the overall HRA analysis.

2.3 Advantages of the Model

The ' model presents two distinct advantages which may make it a preferable dependence modelling technique under certain circumstances.

- The MSF model allows dependent failure probabilities involving multiple human actions to be evaluated on the basis of limited available dependent failure data coupled with structured expert judgment thus, introducing greater objectivity and structure into the assessment of this important element of PRAs. The incorporation of structured expert judgment in dependency is significantly simplified because direct assessment of the level of dependence is not a requirement of the model. Instead, structured expert judgment can be used to estimate a complete set of dependent failure data from which levels of dependence can be calculated. As such, this model allows the results to be less dependent on the subjectivity of an individual analyst, and accordingly, can be expected to improve the consistency of PRA results with regard to dependency analysis.

- The computations involved in calculating the joint failure probabilities from a complete set of dependent failure data can be performed with the FORTRAN computer program provided in Appendix C of this report. The program is user oriented, the input requirements are simple and the output is directly usable in PRA calculations.

2.4 Disadvantages of the Model

The disadvantage of the MSF model at this time, is that the associated uncertainty propagation technique has not yet been developed. As a result, the MSF model provides a single estimate of the multi-component failure probability, but not the associated uncertainty bounds. The optimal use of this model in PRAs requires uncertainty bounds. Research aimed at developing specific means of uncertainty propagation for the MSF model is recommended in Section 5.

3.0 DESCRIPTIVE WALK-THROUGH OF THE MSF MODEL

There are a number of ways to use the MSF model. This section provides a descriptive walk-through of the model and the requirements necessary to use it. A complete discussion and illustrations of using the MSF model are presented in Appendix A.

The process of calculating the multi-component probabilities using the MSF model depends on the type of information available. As discussed in the previous section, the input requirements for the model may be developed either from dependent failure data from actual operating experience or from structured expert judgment. Lack or sparsity of actual data on dependent failures has been the major obstacle to reliable, objective quantification of these failures for use in risk assessments. Accordingly, in most situations, structured expert judgment may be used to establish dependent failure data or complement available actuarial data.

3.1 Input Data to the MSF Model

There are only limited actuarial data on multiple sequential failures in nuclear power plants. If complete data on dependent failures were available (as shown in column 1 of Table 2.1), then calculation of dependence could be easily done using the MSF computer code (Appendix C). However, actuarial data on dependent failures are rare so that some use of partial actuarial data combined with structured expert judgment must normally be used in performing a dependency analysis in a PRA using the MSF model. There are three possible ways to provide input data to the MSF model: (1) using a complete set of actuarial dependent failure data, (2) using a partial set of actuarial dependent failure data combined with structured expert judgment to supplement the available actuarial data, and (3) using only structured expert judgment to generate a complete set of dependent failure data when none exist. Each of these is described below.

3.1.1 Use of a Complete Actuarial Dependent Failure Data

A complete data set of dependent failure data would include the number of times a system was interacted with (i.e., maintained, tested, or calibrated), the number of redundant components in the system, the number of times one component failed, the number of times two (or more) components failed, and the number of times the entire system failed. From these data the independent failure probability of component failure (p) could be calculated and the dependency factor (k) derived by using the MSF computer code. For most systems typically assessed in a PRA, complete data are not available so that structured expert judgment will normally be used.

3.1.2 Integration of a Partial Set of Dependent Failure Data With Structured Expert Judgment

If partial dependent failure data are available, then structured expert judgment can be used to supplement or complete the dependent failure data base which can then be used to calculate the independent component failure probability (p) and the dependence factor (k). For example, if an analysis of Licensee Event Reports (LERs) shows only that there are a certain number of total system failures, structured expert judgment could be used to estimate how often one, two, or three units failed and how often human interactions with the specific system being considered occurred. These data could then be analyzed using the MSF computer code in the same fashion as a complete set of failure data (Section 3.1.1).

3.1.3 Use of Structured Expert Judgment if no Dependent Failure Data are Available

If no dependent failure data are available, structured expert judgment can be used in two ways: (1) to generate a complete set of failure data or (2) to directly estimate the independent failure probability (p) and dependence factor (k).

3.2 Illustration of Using the MSF Model

This section contains an illustration of the use of the MSF model. A complete set of application illustrations is provided in Appendix A.

Consider the case of a Containment Spray Injection System (CSIS) in which there are two redundant recirculation valves. For the CSIS to work, at least one of the two valves must be in the closed position. Both valves are tested monthly by being put into the open position, tested for performance and restored to the closed position following the tests. If the valves are in the open position as a result of a human error (i.e., failure to restore the valves to the closed position after testing), the CSIS is unavailable in the case that it is needed.

If there was no dependence (i.e., complete independence) between the human errors of leaving one valve open and leaving the other open, then the overall system failure probability would be the product of the two separate (i.e., independent) human error probabilities. If the single human error probability for failure to close one valve after testing is 0.01, then the system failure probability due to both valves being open would be 0.0001 (0.01 x 0.01 = 0.0001). However, it is intuitively clear that if an operator testing the system fails (e.g., forgets) to close the first valve, he may be more likely to fail to close the second valve as well. The MSF model can be employed to estimate the actual joint failure probability.

In this illustration, it is assumed that the PRA analyst has available data on failure to close the CSIS valves from the LER file and that the independent probability of failing to restore one CSIS valve is obtainable from available HRA data. These data are presented in Table 3.1 and represent a partial set of input data to the MSF model.

The first step in this analytic process would be to calculate the total number of opportunities for the failures in question to occur. Using the calculation outlined in Appendix A, Section 3.2, the total number of

Table 3.1 Data Available on Failure to Restore CSIS Valves

LER Data:		
Number of times one valve was left open	-	95
Number of times both valves were left open	-	11
HRA Data:		
Accepted probability for failure to restore one valve	-	0.01

opportunities for these errors is 5,270. A complete set of failure data are now available for determination of the degree of dependence and joint failure probability for the CSIS. The complete set of failure data are then input into the MSF computer code which is fully described in Appendix C. The code's output indicates that the joint failure probability for the CSIS failing due to open valves is 0.021. It should be noted that this is a higher value than would be predicted using the assumption of independence (i.e., 0.0001).

If no dependent failure data are available, then structured expert judgment techniques described in NUREG/CRs-3688 and -4016 can be used to develop a complete set of dependent failure data. Alternatively, structured expert judgment can be used to directly estimate the independent failure probability (p) and the degree of dependence (k).

4.0 EVALUATION

In this section the MSF model is evaluated in terms of its practicality, usefulness, and acceptability. As discussed previously, it is a model for use in estimating joint failure probabilities for multiple sequential failures in systems with redundant components.

4.1 Practicality

An evaluation of the practicality of using the MSF model within the context of a PRA must include consideration of the costs, required personnel, and logistics. Each is discussed below.

4.1.1 Costs

The costs of using the MSF model in a PRA are dependent on the number of systems being analyzed for joint failure probabilities due to dependence. Each analysis will require collection of data and/or structured expert judgment. Costs can be broken down as shown in Table 4.1. As indicated in this table, the costs are quite variable. The reason for this is that in some cases complete data will be available which would result in low costs since no structured expert judgment is necessary and the data analysis is handled by the MSF computer code. On the other hand, when no data are available, it would take less time to determine the unavailability of the data, but the structured expert judgment would require more resources. As a result, it can be concluded that use of the MSF model is most practical, on a cost basis, if a complete set of failure data are available. When no data are available, the MSF model will entail higher costs.

It must be noted that a well planned PRA process would involve predetermination of which systems do not have complete data available and aggregation of these analyses to take advantage of economies by having the expert group meet to address all of the systems in one session or set of sessions. For the purposes of assessing the costs of using the model in a PRA, assume

Table 4.1 Costs of Using the MSF Model for a Single System Analysis

	Complete Actuarial Data Available (Staff-Days)	Partial Data and Structured Expert Judgment (Staff-Days)	Structured Expert Judgment and No Data (Staff-Days)
Data Collection	1	2	1
Structured Expert Judgment*	0	4	7
Data Analysis	0.5	1	2
Total	1.5	7	12

*Assumes a three-member expert group and one facilitator.

that ten systems are involved that must be analyzed in terms of human induced dependence. Of these ten systems, assume no data are available for six of them, partial data are available for three of them, and complete data are available for only one. Assuming further, that expert groups address all of these systems together (taking 6 days), the costs which would result are presented in Table 4.2. It is concluded that the estimated cost in Table 4.2 (44 staff-days) is not prohibitive considering the critical importance of these analyses in a PRA and is, therefore, practical on a cost basis.

Table 4.2 Cost for 10 Redundant Systems Analyses

Data Collection	7 Staff-Days
Structured Expert Judgment*	24 Staff-Days
Data Analysis	13 Staff-Days
Total	44 Staff-Days

*Assumes a three-member expert group and one facilitator.

4.1.2 Personnel

A PRA specialist may be capable of analyzing available actuarial data (e.g., the LER file) to develop a complete set of data on events associated with dependent failures in redundant systems and of inputting the data into the MSF computer code. However, it will be necessary to have a qualified

human reliability professional develop any data which are not otherwise available. For instance, if any use of structured expert judgments are made, an individual qualified in facilitating the group decision process must be employed. In addition, any expert group should be comprised of a mix of experts in operations, human reliability, and risk assessment. The minimum expert group size should include three experts and one facilitator.

4.1.3 Logistics

In order to optimize the use of logistical support in estimating the joint failure probabilities of components in redundant systems, it would be desirable to consider all dependence calculations used in the PRA in one session or set of sessions. This would allow for use of the same expert group throughout the dependence estimation process. As a result, the MSF model should be initiated after critical systems employing redundancy are identified.

4.2 Usefulness

An evaluation of the usefulness of the MSF model in the PRA process must include consideration of the justification for its use, the compatibility and comprehensiveness of the model with regard to PRA process, and the validity of the model as a human reliability analysis tool. Each is discussed below.

4.2.1 Justification

Dependency analyses play a very critical role in the overall assessment of risks from plant operation. Accordingly, the development of a new dependency model is justified because it provides PRA practitioners with a new method to use which may be considered preferable to existing methods (e.g., NUREG/CR-1278, 1983; WASH-1400, 1975) which have been the subject of criticism (NUREG/CR-0400, 1978).

The MSF model offers a more rigorous framework for estimating the level of dependence and the joint failure probabilities (i.e., total system failures) that result. Thus, it is less dependent on the subjective analysis of a individual HRA specialist as is currently done using other available models.

4.2.2 Compatibility With the PRA Process

Use of the MSF modeling adds more consistency to the estimates of dependence used in PRAs. If used in the optimal logistic fashion, the dependency estimation process using the MSF model would be done at the same time for all human induced dependence calculations considered in the PRA. The same expert group should be used to make estimates of dependent failure data, independent failure probabilities, and/or dependency levels. As a result of using a single expert group, there should be improved consistency in making these estimates.

The output of the MSF model is in the form of system failure probabilities attributable to human interactions. This form of output is fully compatible with the needs of PRA.

4.2.3 Validity of the Model

The estimation (i.e., prediction) of the level of dependence introduced by human interactions with safety-related systems is a very important factor in the overall quality of any PRA. In order to be conservative in PRAs, human induced dependence is assumed to be a factor included in any calculation of a joint failure probability for redundant components in a system. Various models, discussed in detail in Appendix B, Section 1.1, have been used in PRAs, but none had been tested using a complete set of actual human performance data generated for that purpose. The central question in testing the model is whether, in sequential human task situations where dependence may exist, the model can predict a joint failure probability in a manner significantly more accurate than would be predicted by chance. In essence, the

question is whether the MSF model is a useful tool in accounting for how human beings actually commit sequential errors. If the MSF model fits actual data on humans committing errors during sequential tasks, it can be useful in a PRA.

In order to assess the validity of the model, this section describes two efforts: (1) an internal validity check and (2) an external validity check. Each is described below.

4.2.3.1 Internal Validity Check

In order for a model to be used, it must perform as expected in terms of the internal validity of inputs and outputs. Sensitivity analysis was used to address the question of validity of the model. In sensitivity analysis, predefined inputs are provided and results obtained from the model are compared with expected outputs to assess whether the interpretation of input data by the model is appropriate. Accordingly, sensitivity analyses were designed with several different types of input data on dependence. The results of the sensitivity analysis provide the following observations indicating the internal validity of the model.

- For cases with representing random double, triple, or quadruple failures, (i.e., no dependence), the model consistently yields dependency factors (k) which are extremely close to zero.

- As the data input are modified to represent an increase in dependence (i.e., increasing sequential error rates), the model calculates higher values of the dependency factor (k), and the joint failure probability is appropriately increased.

4.2.3.2 External Validity Check

In order to reliably test the external validity MSF model, a small-scale simul tion experiment was undertaken which is described in detail in Appendix

B. Two previous attempts to test the external validity of the MSF model using actuarial data (NUREG/CR-2211, 1981; ASA-709, 1982) were constrained by the limited human performance data on sequential errors which was available. In order to collect sufficient human performance data, three analogues to maintenance, testing, and calibration tasks, were developed using actual nuclear power plant procedures of several representative tasks and reviewed by individuals with appropriate nuclear power plant operational experience. Work characteristics which were simulated in the experimental tasks were: (1) routine, repetitive performance, (2) intermittent scheduling interspersed with other tasks, (3) well-practiced subjects with relatively low error rates, (4) no direct feedback on errors and no direct supervision of work, (5) continuously available written procedures for each task, and (6) the option of restarting a task without penalty.

Five paid, male subjects were selected and given extensive training on the three tasks. After training was completed, subjects worked four hours per day, five days (i.e., 20 hours per week) performing these tasks. Usable data from eight to eleven sessions were collected for each subject.

After the data were collected, they were statistically analyzed (i.e., chi-square) to determine if, when sequential errors occurred, the MSF model accounted for the sequential error rates evidenced by the actual performance of the subjects to a statistically significant level. For sequential errors found in the data, the MSF model fit the patterns of how actual error probabilities increased to a significant level (i.e., significant to $p < 0.10$). The fit was sufficiently good that the MSF model reliably described the performance of the subjects on sequential tasks.

4.3 Acceptability

In light of the successful test of the MSF model in the small-scale psychological experiment described in the previous section and detailed in Appendix B, the MSF model has credibility as a tool in PRA. The MSF model was also reviewed by several PRA practitioners who have indicated a willingness to

use the model if it is documented for use. Because this report provides a users' guide (Appendix A) and the MSF computer program (Appendix C), it is now usable within the PRA community.

5.0 SUMMARY, CONCLUSIONS, AND RECOMMENDATIONS

5.1 Summary

This report presents an evaluation of the MSF model in terms of use in the PRA process. It presents an overview of the model (originally described in NUREG/CR-2211, 1981) and provides an evaluation of its practicality, usefulness, and acceptability, along with guidance and instructions on using the model (Appendix A).

5.2 Conclusions

The estimation of multiple failure incorporating human dependency has so far been subjective and the model presented provides an improvement of the state-of-the-art in terms of reducing the level of subjectivity in the analysis. The model, in an ideal situation requires data on dependency which may, in many cases, be currently available. However, the model provides a means for using any limited data that are available and provides a systematic means of incorporating structured expert judgments to support them.

5.3 Recommendations for Future Work

- Various dependency situations due to human interactions that are encountered in performing a PRA should be categorized and classified into generic groups based on tne analysis of the dependency conditions. Attempts should then be made to develop data bases representing the various situations so that the MSF model can be used to develop dependent failure probabilities which may then be used directly in PRAs.

- Modeling of human error dependency is rather complex and, accordingly, the MSF model requires a number of assumptions. In light of the present state of knowledge concerning dependence and risk assessment, these assumptions are considered appropriately

conservative. However, with the accumulation of data on dependent failures, these assumptions should be examined and modifications in the model should be sought if appropriate.

- The MSF model does not provide any means of obtaining uncertainty bounds in the estimation of dependent failure probabilities. Statistical procedures are available for developing uncertainty bounds (NUREG/CR-2211, 1981) and efforts should be initiated in that direction.

- NRC should initiate efforts to modify its event and incident reporting systems to collect dependent failure data in order to provide input to this important area of risk assessment.

REFERENCES

ASA-709, "Conceptual Field Validation of the Models of Multiple Sequential Failures During Testing, Maintenance and Calibration," D. L. Schurman and J. K. Hawley, Applied Science Associates, Inc., August 1982.

Epler, E. P., "Common-Mode Failure Considerations in the Design of Systems for Protection and Control," Nuclear Safety, 10(1), January-February 1969.

Fleming, K. N., et al., "HTGR Accident Initiation and Progression Analysis Status Report: Phase II Risk Assessment," U.S. Department of Energy, Washington, DC, 1978.

Fleming, K. N., Mosleh, A., and Kelley, A. P., Jr., "On the Analysis of Dependent Failures in Risk Assessment and Reliability Evaluation," Nuclear Safety, 24(5), September-October 1983.

NUREG/CR-0400, "Risk Assessment Review Group Report to the U.S. Nuclear Regulatory Commission," H. W. Lewis, et al., 1978.

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

NUREG/CR-2211, "Modeling of Multiple Sequential Failure Probabilities During Testing, Maintenance, and Calibration," P. N. Samanta and S. P. Mitra, Brookhaven National Laboratory, 1981.

NUREG/CR-2300, "PRA Procedures Guide," Volumes 1 and 2, American Nuclear Society and Institute of Electrical and Electronics Engineers, 1982.

NUREG/CR-2782, "Interim Reliability Evaluation Program (IREP) Procedures Guide," D. D. Carlson, Sandia National Laboratories, 1982.

NUREG/CR-2815, "Probabilistic Safety Analysis Procedures Guide," I. A. Papazoglou, et al., Brookhaven National Laboratory, (in press) 1985.

NUREG/CR-3010, "Post Event Human Decision Errors: Operator Action Tree/Time Reliability Correlation," R. E. Hall, et al., Brookhaven National Laboratory, 1982.

NUREG/CR-3519, "Human Error Probability Estimation Using Licensee Event Reports," K. J. Voska and J. N. O'Brien, Brookhaven National Laboratory, 1984.

NUREG/CR-3688, "Generating Human Reliability Estimates Using Expert Judgment," M. K. Seaver, et al., Sandia National Laboratories, 1984.

NUREG/CR-4016, "Application of SLIM-MAUD: A Test of an Interactive Computer-Based Method of Organizing Expert Assessment of Human Performance and Reliability," E. A. Rosa, et al., Brookhaven National Laboratory, (in press) 1985.

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, 1975.

APPENDIX A

OVERVIEW OF THE MODEL AND A USERS' GUIDE

1.0 OVERVIEW OF THE MODEL

1.1 Basic Philosophy

The multiple sequential failure (MSF) model provides a statistical framework for quantifying the probability of failure on two or more tasks performed by one or more persons. The method uses the available limited data and the understanding of the task situation to derive the joint probability of failure. Since the model attempts to derive the joint failure probability based on the failure data, it observes the effect on the system, i.e., the number of times different groups of components affected, without attempting to distinguish the cause in terms of number of persons involved in the task. Thus, the observed data are considered to embody the dependence between and within the people, if more than one person is involved in carrying out the task.

The MSF model embodies various characteristics of dependent human failures to develop the analytical framework for estimating the available data. Without the incorporation of such characteristics, the estimation of joint failure probability will require substantial amount of data; which is not available. Otherwise, the estimation will be grossly erroneous.

Whenever an individual or a group of people are involved in more than one task at a time, there is a level of dependence from one task to another. These tasks are usually similar and repetitive in nature. For example, during the monthly test of auxiliary feedwater systems, three pairs of pumps discharge valves are closed and they should be left open following the test. If the operator fails to open the first valve following the test, the probability of failure to open the second valve is dependent on the first act, and similarily, the probability of failure to open the third valve is dependent on the first two acts The estimation of the degree of dependence among the tasks, and thus, the avoidance of underestimation in the joint failure probability is of primary importance in reliability calculations.

An analysis of the task situations where humans are involved in carrying out multiple tasks reveal the sequential or cascading nature of failure. Since humans perform their task in a sequence, and the procedures demand the observance of such sequences, the failures propagate in one direction only. For example, consider the failure to open three valves, A, B, and C by an operator. The procedure demands him to open valve A, then valve B, and then valve C. Accordingly, the failure to open valve A will influence his probability of failure to open valves B and C. But his failure to open valve B, influences his probability of failure to open valve C and not the probability of failure to open valve A; since if he already performed his task on valve A, he has no reason to go back to A, and if he does go back, the whole process starts over again. Thus, a failure

in valve B can result in failure of two of the components and failure in valve C will result in a single failure.

This analysis of dependence among the failures provides the framework for the probabilistic modeling of multiple sequential failure. Here, the degree of dependence is considered a continuum, and a positive dependence is assumed among the failures. This implies a positive dependence between events, i.e., a failure on the first task implies an increased probability of failure on the second task. The modeling of this phenomenon is accomplished by defining the range within which the probability of second failure lies. Given two sequential actions, the probability of failure in the second action given failure in the first action is increased from the independent failure probability of the second action, and thus, must lie between the lower bound of the independent failure probability and the upper bound of 1. The increased failure probability is considered to a fraction of the entire range and is estimated from data. When more than two actions are involved, the degree of dependence is expected to be of more complicated in nature. Given three sequential actions, the probability of failure in the third action, given failure in the first two, should exceed its independent failure probability by an amount greater than that in the case of second failure, given the first failure. In this model, it is argued that in multiple sequential failures the probability of failure in the third action is more dependent on the probability of failure in the second action, given the first failure, than on the independent failure probability of the third action. Thus, the range of the probability of failure in the third action is shrunk by increasing its lower bound from its independent failure probability to the probability of second failure, given the first failure. For similar dependent actions, the probability of second failure, given the first failure is always greater than the independent failure probability of the third action. Thus, the range of the probability of failure in the third action is shrunk by increasing its lower bound from the independent failure probability to the probability of second failure, given the first failure. In this way, the ranges of probability of failure of higher and higher actions are shrunk, and the respective probabilities are increased because of the change in the lower bound of the respective shrunken ranges. To some extent, this approach authomatically takes into account the increased dependence for any action compared with the preceeding action in a multiple sequential action. That is, such increased dependence is built into the model.

1.2 Assumptions

Development of a mathematical model of any real situations is an idealization of various underlying processes and implicit in it are a number of assumptions. Such assumptions are necessary ingredients of the model development and in no way undermines the model as long as these are consistent with our understanding of the process. In fact, these assumptions allow one to develop the analytic model of the complex structure. In this case, many of the assumptions are derivative of the basic philosophy behind the model and others are due to lack of detailed data on dependent actions. As more and more data become available, number of assumptions can be removed; however, in the meantime, care is taken to insure that the impact of these assumptions are conservative.

The following provides a summary of the various assumptions implicit and explicit in the model:

1. Dependent human failures progresses in a sequential manner and follows a defined direction. This assumption directly relates to the way tasks are performed and the procedures on the task.

2. Positive dependence was modeled between successive failures. Negative dependence, implying that the failure on the first task may decrease the probability of failure in the second task, was not taken into consideration. There may be such instances; however, the evidence is not clear and this assumption is expected to introduce conservativeness in the estimate of the joint failure probability.

3. No dependences on successes were assumed. If there is positive dependence on successes, the neglect of such dependencies will add to the conservativeness in the model. But, if negative dependence is present, this is expected to result in underestimation. The relative effect of positive and negative dependence on successes is not known and the overall impact of this assumption cannot be judged at this point. None of the other dependence models accounts for the success dependencies at this time.

4. The conditional probability failure given success is assumed to be the independent failure probability of that action. For example, in the case of three tasks, if there is failure in the first task and success in the second task, then the probability of failure in the third task is assumed to be independent of the previous tasks. This assumption is plausible because, as soon as a correct action is performed following a failure, the dependence on that failure is assumed to be lost.

1.3 Comparison with Related Model

The dependence model presently used in assessing human error dependence in PRAs was developed by Swain and Guttmann (NUREG/CR-1278, 1983). Previous discussions in this report at various points provided contrasting features of the MSF model with that of Swain and Guttmann. In this section, a brief summary of major differing points are provided.

1. The dependence failure modeling in NUREG/CR-1278 is based on subjective assumption and there are no means of incorporating actual data, even if they are available. At present, some actual data are available and the MSF model provides a vehicle for improving the state-of-the-art by removing subjectivity in the analysis.

2. In NUREG/CR-1278, the analyst is restricted to choosing amongst the specified levels of dependence. There is no such restriction in the MSF model. The degree of dependence is assumed to be a continuum as it correctly should be.

3. The dependency treatment is NUREG/CR-1278 is arbitrary and highly questionable. The conditional failure probability on Task N in Swain and Guttmann's dependence modeling given failure in the immediately preceding Task N-1for a given level of dependencce is given by:

$$Pr[F_{"N"}|F_{"N-1"}|LD] = Pr|F_{"N"}] + 1/2\ (1-Pr[F_{"N"}])$$

where LD represents the low level of dependence, the interesting point to note that the conditionality of failure in Task N (i.e., the failure in Task N-1) is completely absent from the right hand side of the equation which calculates the conditional failure probability. The conditional failure probability is thus dependent on the present task which is contrary to the basic process in dependent failure. The MSF model incorporates the dependency phenomenon in it by analyzing the task situation involved.

4. The dependence modeling in NUREG/CR-1278 only addresses 1 out of n:G logic-type situation and the method is not capable of handling more general n out of n:G logic-type configuration. Various logic configurations are encountered frequently in PRA for nuclear power plants and the analyst is left without any guidance for evaluating these situations. The MSF model addresses all types of situations commonly encountered in PRAs and provides specific formulas for use in each of these situations.

1.4 Limitations and Area of Application

The MSF model is based on human error task analysis that the dependency is cascading in nature as explained in Figure 2.1. It is believed that dependency in human errors is cascading in nature, however, if there is reason to believe that the nature of dependency is different, the MSF model is not the appropriate approach.

The MSF model can be used when actual data or data based on structured expert opinion are available. It can also be used when the degree of dependence is directly assessed based on structured judgments. The model is capable of handling different levels of degree of dependence between different pairs of actions. However, the equations provided in this report relates to situations where the degree of dependence is assumed to be constant across tasks. Section 3 of the report provides guidance on handling situations where the degree of dependence is different within subsets of tasks.

2.0 MULTIPLE SEQUENTIAL FAILURE MODEL

2.2 Governing Equations

Let us consider n sequential actions and let the symbols H_1, H_2, .. H_n represent failure in actions, respectively. Correspondingly, $\bar{H}_1$, $\bar{H}_2$,...$\bar{H}_n$ represent success on the actions. The expression $H_1H_2H_m$ represents repetitive failures in action 1 through m.

$P(H_i)$: Probability of failure on the ith action
$P(H_1H_2...H_n)$: Joint probability of n sequential failures.

Considering dependencies, the joint probability can be written as:

$$P(H_1H_2,...H_n) = P(H_1)\ P(H_2/H_1)\ P(H_3/H_1H_2)\ ...\ P(H_n/H_1H_2...H_{n-1})$$

where the bounds of the conditional probabilities are given by:

$$\begin{array}{c} P(H_2) \leq P(H_2/H_2) \leq 1 \\ \mathrm{Max}\ [P(H_2/H_1), P(H_3)] \leq P(H_3/H_1,\ H_2) \leq 1 \\ \vdots \qquad\qquad \vdots \qquad\qquad \vdots \\ \mathrm{Max}[P(H_{n-1}/H_1H_2...H_{n-2}), P(H_n)] \leq P(H_n/H_1, H_2...H_{n-1}) \leq 1 \end{array} \qquad (1)$$

In most of the practical situations of importance in nuclear power plant applications, the dependency is considered among similar actions and accordingly, the failure probability of the indepedent actions are the same. However, the model is not limited to similar action but the remainder of the discussion, for ease of understanding, will be carried on within that restriction. Mathematically, this implies:

$$P(H_1) = P(H_2) = ... = P(H_n) = p$$

2.1.1 Calculation of Conditional Probabilities

The determination of conditional probability is the most difficult part in the dependency modeling and contains the essential features of the MSF model. Here, the conditional probability of failure in the second action, given failure in the first action is expressed as the sum of the independent failure probability (p) and a dependent failure probability (P_{df}).

$$P(H_2/H_1) = p + p_{df}$$

The dependent failure probability describes the increase in the failure probability due to dependency and is expressed as a fraction of the total range of the conditional failure probability.

$$p_{df} = (1-p)k \tag{2}$$

where:

$(1-p)$ = the total range of $P(H_2/H_1)$ from Equation (1)
k = the dependence factor.

Accordingly:

$$P(H_2/H_1) = p + (1-p)k$$

The probability of failure in the third action, given failure in the first two, is given by:

$$\begin{aligned} P(H_3/H_1H_2) &= p + (1 - P(H_2/H_1))k \\ &= p + (1-p)\ [1-(1-k)^2] \end{aligned}$$

where the dependence factor is assumed to be the same as in the previous case. In a general situation, k's can be different.

Proceeding in a similar manner, one can show:

$$P(H_n/H_1H_2 \ldots H_{n-1}) = p + (1-p)[1-(1-k)^{n-1}] \tag{3}$$

Other types of conditionality appear in the development of joint failure probability. The probability of failure following a success is considered to be the independent failure probability.

$$P(H_2/\overline{H}_1) = P(H_2) = p$$

Also, the probability of failure of the ith action, given success in the (i-1)th action is considered to be the independent failure probability of the ith action, irrespective of the failure history in the 1st to (i-2)th action (assumption 4, see Section 2.2). That is, for three actions, the probability of failure in the third action, given succeess in second and failure in the first is given by:

$$P(H_3/\overline{H}_2H_1) = P(H_3) = p$$

2.2 Users' Guide for Applications to Different Task Situations

2.2.1 2-Unit System

For a 2-unit system, the system logic configuration of interest is 1 out of 2:G logic, i.e., the system is good if at least one of the two units is good.

The failure probability (${}^{1}P_2$) for 1 out of 2:G logic is given by the MSF mode as:

$$1 \text{ out of } 2{:}G \text{ logic} = {}^{1}P_2 = p^2 - kp^2 + kp \ .$$

The moments equations to be solved for estimating p and k from data are:

$$2p + kp(1-p) = 1/N\ (x_1+2x_2)$$

$$2p + 2p^2 + 3kp(1-p) = 1/N\ (x_1+4x_2) \ .$$

where:

N = total number of opportunities
x_1 = number of times one unit is failed
x_2 = number of times both the units are failed.

2.2.2 3-Unit System

For a 3-unit system, the system logic configurations of interest are 1 out of 3:G logic and 2 out of 3:G logic.

The failure probabilities for these logic configurations given by the MSF model are as follows:

$$1 \text{ out of } 3{:}G \text{ logic} = {}^{1}P_3 = \prod_{i=0}^{2} \left[1-(1-p)(1-k)^i\right]$$

$$2 \text{ out of } 3{:}G \text{ logic} = {}^{2}P_3 = p(2-p)-2p(1-p)^2(1-k)$$

The moments equations to be solved for estimating p and k from data are ($p<<1$).

$$3p + 2pk + 2k^2p = 1/N\ (x_1 + 2x_2 + 3x_3)$$

$$3p + 6p^2 + 6pk + 10k^2p = 1/N\ (x_1 + 4x_2 + 9x_3)$$

where:

N = total number of opportunities
x_i = number of times i units have failed.

2.2.3 4-Unit System

For a 4-unit system, the system logic configuration of interest are 1 out of 4:G logic, 2 out of 4:G logic, and 3 out of 4:G logic.

The failure probabilities given by the MSF model are as follows:

$$1 \text{ out of } 4{:}G = {}^1P_4 = \prod_{i=0}^{3} [1-(1-p)(1-k)^i]$$

$$2 \text{ out of } 4{:}G = {}^2P_4 = p(2-p) - 2p(1-p)^2(1-k) + p(1-p)(p^2+p-2)(1-k)^2 + 2p(1-p)^3(1-k)^3$$

$$3 \text{ out of } 4{:}G = {}^3P_4 = (3-3p+p^2)\, p-3p\, (1-p)^3(1-k)$$

The moments equations to be solved for estimating p and k are ($p<<1$).

$$P(1/4) + 2P(2/4) + 3P(3/4) + 4P(4/4) = 1/N\, (x_1 + 2x_2 + 3x_3 + 4x_4)$$

$$P(1/4) + 4P(2/4) + 9P(3/4) + 16P(4/4) = 1/N\, (x_1 + 4x_2 + 9x_3 + 16x_4)$$

where:

$P(1/4) = 4P - 12p_2 - 3pk$,
$P(2/4) = 3pk + 6p^2 - 4pk^2$,
$P(3/4) = 8p^2k + (4p-26p^2)k^2 - (8p-37p^2)k^3 + 9pk^4 - 5pk^5$,
$P(4/4) = 13p^2k^2 + (6p-31p^2)k^3 - 9pk^4 + 6pk^5$,
N = total number of opportunities, and
x_i = number of times i units are failed.

3.0 ILLUSTRATIONS OF USING THE MSF MODEL

Task processing, or calculating the dependent failure probability using the MSF model depends on the type of information available. As discussed in the preceding section, the input requirement for the model may be developed either from actual operating experiences or from structured expert judgments. Lack or sparsity of actual data on dependent failures has been the major obstacle to its quantification. Accordingly, in many situations structured expert judgment may be used to establish or complement the data base used as input to the MSF model.

The adequacy of dependent failure data base has two aspects. One, dependent failures are rare events and need to be quantified in situations which are reliable to start with. Accordingly, the number of dependent failures to be observed is expected to be small unless data are gathered over a very long period of time. The other aspect relates to the situation where no attempt has been made to collect the experiences of dependent failures. In the first instance, if the data base is developed from LERs, the number of opportunities are rather difficult to estimate, even if the number of dependent failures are gathered. The inadequacy of the first type of data is expected to be the norm of dependent failure analysis and, accordingly, the MSF model was developed to derive the failure probability in terms of limited data and structured expert judgment along with a mathematical model. In the second instance, structured expert opinion has to be used to develop a data base or to develop a level or degree of dependence among the actions considered.

The actual data on dependent failures may be available in two different forms:

1. A complete set of dependent failure data (i.e., the total number of opportunities and the number of failures of each combination of components are available). For example, in a 3-unit system, the data on the total number of opportunities along with the number of observations representing each failure in one of the units, two of the units, and all three of the units failed are available.

2. A partial set of dependent failure data (i.e., only the number of failures of each combination of components are available). For example, in a 3-unit system, only the number of observations representing each failure in one of the units, two of the units, and all three of the units are available.

In situations where actual data are not available, structured expert judgment can be used in one of the following ways:

1. Develop a complete set of dependent failure data (i.e., on the basis of the analysis of the tasks using structured expert judgment will provide the expected number of failures for each combination of components for a given number of opportunities).

2. Develop partial set of dependent failure data (i.e., given the number of single failures, how many times other combinations of failures are expected to be developed from structured expert opinion). That is, given 10 single failures for a 3-unit system, how many times does one expect to observe failure of two of the units or of all three of the units?

3. Directly estimate the degree of dependence (i.e., structured expert judgment is used to directly estimate k from an analysis of the task situation). The use of structured expert judgment at this level is most difficult and introduces maximum uncertainty because of lack of understanding of the nature of dependency in general and the necessity of completely understanding the MSF model to establish a parameter k directly in the model.

From the above discussion, the types of inputs that may be available for quantifying multiple sequential failures probabilities can be categorized as follows:

1. Complete data on dependence are available. The complete data on dependence (see, for example, Table 1.1) has been either obtained from actual operating data or developed from structured expert judgment. In this case, the MSF model requires no other input and the model will provide the failure probabilities for various logic configurations.

2. Partial data on dependence are available. Partial data on dependence, as defined earlier, are either available from actual operating data or can be developed using structured expert judgment. The MSF model in this case uses the available independent failure probabilities (p) to develop the total number of opportunities.

3. Degree of dependence (k) is available. Because actual data on dependence are lacking, structured expert judgments are used in this situation to develop the degree of dependence (k) among the tasks. In this case, the MSF model uses the independent failure probability (p) and the k developed from structured expert judgment to determine the multiple sequential failure probabilities.

In evaluating multiple sequential failure probabilities for multi-unit systems, two situations arise:

Situation 1. The degrees of dependence between the tasks are similar (i.e., in a 3-unit system where the human errors occurring in each of the units are similar, the degree of dependence between units 1 and 2 is expected to be similar to that between units 2 and 3).

Situation 2. The degrees of dependence within a subset of tasks are different (i.e., in a two-train system each consisting of two redundant components, the degree of dependence between the redundant components may differ from that between the trains). In addition, the degree of

dependence between two redundant components can be different from one train to another.

The types of inputs available in each of the two above situations can be any of the three previously discussed categories or a combination of them.

In the following sections, for several typical scenarios, the steps involved in using the MSF model are provided with example illustrations. According to different dependence situations and the type of data inputs available, four scenarios expected to be commonly encountered in quantifying human error dependency in testing, maintenance, and calibrations are considered.

1. Dependencies between different units of a multi-unit system are similar and the degree of dependence is estimated directly from structured expert judgments (see Section 3.1).

2. Dependencies between different units of a multi-unit system are similar but partial data on dependency are available (see Section 3.2).

3. Dependencies between different units of a multi-unit system are similar and complete data on dependency are available (see Section 3.3).

4. Dependencies within a subset of tasks are different and the degree of dependence is estimated from structured expert judgment (see Section 3.4).

The computations used in the MSF model can be performed efficiently with the FORTRAN computer code presented in Appendix C, and this code is appropriately referred to in the task steps.

3.1 Dependencies Between Different Units Are Similar and the Degree of Dependence (k) Is Estimated Directly from Structured Judgments

This situation applies if the analyst has no data on dependent failure; however, the independent failure probability is available. Because of lack of data, the degree of dependence is estimated on the basis of structured expert judgment. The level of dependence derived on the basis of the subjective judgment can be used at this level in the MSF model to obtain the dependent failure probabilities. However, in the case of the MSF model, because the dependence is assumed to be a continuum, no restriction exists on the level of dependence to be estimated, as in the case of other model (e.g., NUREG/CR-1278, 1983).

3.1.1 Example Calculation

Consider the situation of estimating dependent failure probability based on the possibility of repetitive human errors during containment pressure sensor comparator calibration and test procedure. In this case, a set of four comparators, all similar in appearance and function, are tested monthly, and the calibration, test, and adjustment procedures for all comparators in the set are identical. It is likely that an incorrect action performed on one unit would be

repeated on other units of the same type where a particular set of actions are called for in the procedures. Therefore, the miscalibrations or damaging of all four of the pressure comparators are dependent. Also, the system configuration is such that failure of any two of the four pressure transducers will fail both the trains in the system. In performing the probabilistic risk assessments, the dependent failure probability for this situation needs to be determined in quantifying the fault trees associated with the system for which the comparators are a part. Consider also in this situation that no data on dependent failures are available. Following the flow diagram in Figure A.1, we present a step-by-step illustration of how to calculate the dependent failure probability using the multiple sequential failure model.

<u>Step 1 - Determine Number of Dependent Actions Involved (r)</u>

In this step we define the total number of dependent actions involved in the task. In our example problem, since the calibration is performed on a set of four comparators, the total number of dependent actions is four, (i.e., $r = 4$).

<u>Step 2 - Define System Success Configuration</u>

Once the total number of dependent actions or units have been defined, the system success configuration must be defined. This means that the system analyst must understand how many of these actions will cause the system to fail. In our example problem, failure in any two of the four pressure comparators will fail the system. Thus, the system failure configuration is 2 out of 4, which information can be transformed into success configuration, commonly known as the system G-logic configuration[2]:

$$G = r - F + 1 ,$$

where,

F = the failure configuration,
r = the total number of units.

In the example, $G = 4 - 2 + 1 = 3$ (i.e., 3 out of 4:G logic system). This means that the system is good if and only if the actions in 3 of the units are successful. In many situations, a multi-unit redundant system is directly defined in G-logic configuration, and one does not need to perform the transformation.

[2] k out of n:G logic configuration signifies that the system of n components is good if f at least k components are good.

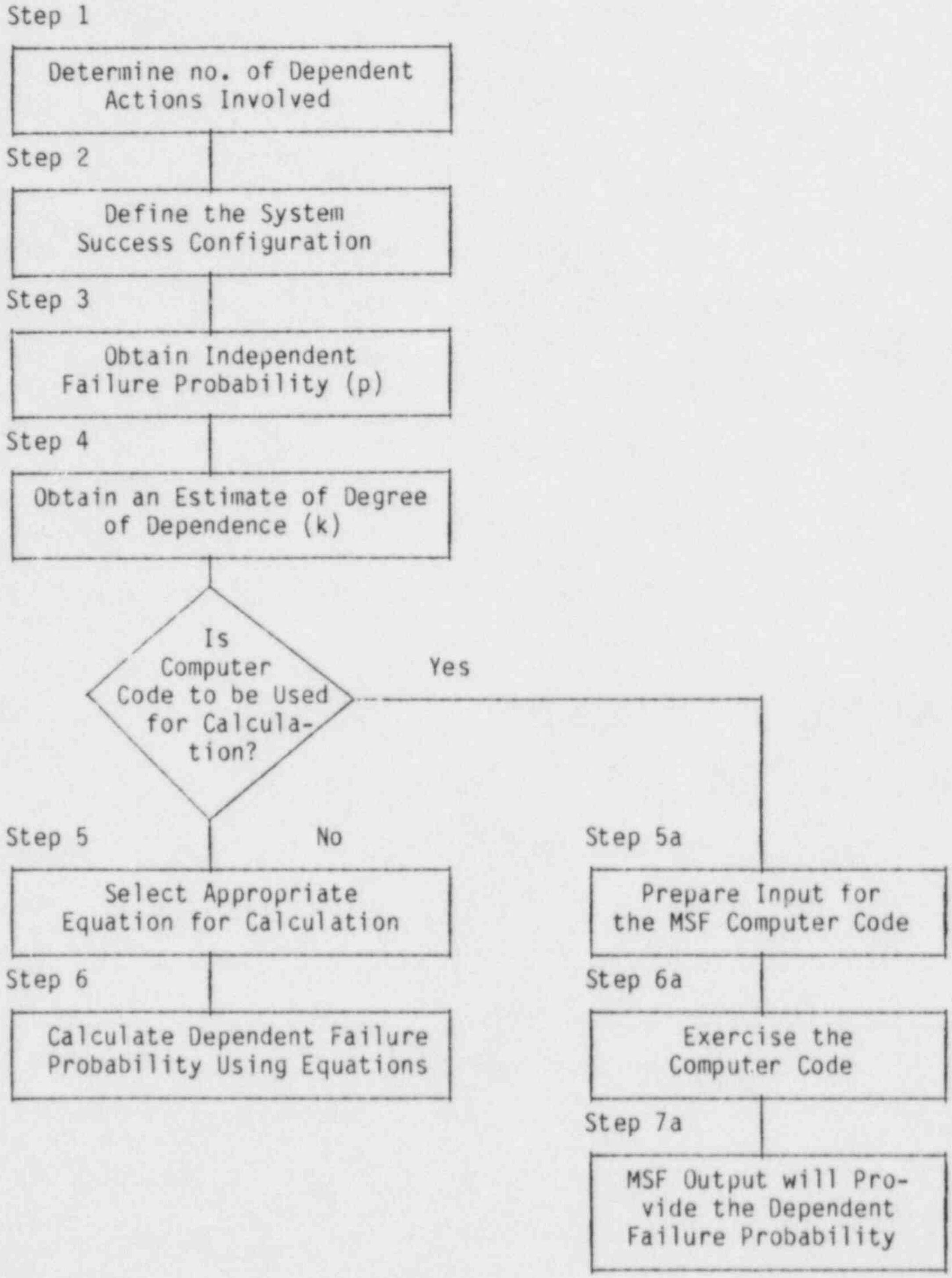


Figure A-1 Flow diagram when p and k are known.

Step 3 - Obtain Independent Failure Probability (p)

Independent failure probability (p) (i.e., the failure probability of the action if it had been performed separately, under the same condition) is assumed to be available. In the example problem, the independent failure probability of miscalibrating a pressure comparator is assumed to be 1×10^{-2} (i.e., $p = 1 \times 10^{-2}$).

Step 4 - Obtain an Estimate of Degree of Dependence (k)

In this case, since no data are available on dependent failures, the degree of dependence among the actions is to be based on structured expert judgments. In the MSF model, the degree of dependence is assumed to be a continuum and an analyst can choose any value between its appropriate boundary of 0 and 1 based on the analysis. In our example problem, let us assume that k is assigned a value of 0.1 for the situation described.

Step 5 - Selection of an Appropriate Equation for Calculation

With the information gathered in Steps 1 to 4, this step involves selection of the appropriate equation for the calculation of dependent failure probability if it is decided not to use the MSF computer code. For this situation, the calculation is rather straightforward and can easily be performed on a hand calculator.

For the example problem, the reliability configuration is 3 out of 4:G logic, and the corresponding dependent failure probability (${}^3P_4$) equation given by the MSF model (see Table A-1) is

$$ {}^3P_4 = (3-3p + p^2)\ p - 3p\ (1-p)^3\ (1-k)\ . $$

Step 5a - Prepare Input for MSF Computer Code

If the user decides to use the MSF computer code for performing the calculation, the input to the code should be prepared at this point. Instructions for preparing the input to the MSF code are given in Appendix C.

Step 6 - Calculate Dependent Failure Probability

At this point, the equation for calculating dependent failure probability has been determined and, also, the unknowns p and k are known. Accordingly, the dependent failure probability can be calculated.

For the example problem, $p = 1 \times 10^{-2}$, $k = 0.1$, and

$$ \begin{aligned} {}^3P_4 &= (3-3p + p^2)\ p - 3p\ (1-p)^2\ (1-k) \\ &= 3.2 \times 10^{-3} \end{aligned} $$

Thus, the probability that the operator will fail at least two of the four pressure comparators during calibration and testing is 3.2×10^{-3}.

Table A-1 Multiple Failure Probabilities Due to Human Error for Different G-logic Types of Systems.

Type of G-Logic System	System Failure Probability
1 out of 2	$kp + p^2(1 - k)$
1 out of 3	$p - p(1 - p)(1 - k) - p(1 - p)(1 - k)^2 + p(1 - p)^2(1 - k)^3$
2 out of 3	$p(2 - p) - 2p(1 - p)^2(1 - k)$
1 out of 4	$\prod_{i=0}^{3} [1 - (1 - p)(1 - k)^i]$
2 out of 4	$p(2 - p) - 2p(1 - p)^2(1 - k) + p(1 - p)(p^2$ [illegible] $p - 2)(1 - k)^2 + 2p(1 - p)^3(1 - k)^3$
3 out of 4	$(3 - 3p + p^2)p - 3p(1 - p)^3(1 - k)$

Step 6a - Exercise the Computer Code

This step required the exercising of the computer code with the input deck for the MSF computer code prepared in Task 5a.

Step 7a - Results from the MSF Computer Code Output

Once the computer code has been successfully run, the output will provide the desired dependent failure probability. In the example problem, the dependent failure probability generated by the computer code is 3.2×10^{-3}.

3.2 Dependencies Between Different Units Are Similar and Partial Data on Dependency Failures are Available

Reliability analysts often find that partial data on dependence are available and it is to their advantage to use this information to obtain an estimate of the dependent failure probability. In any quantification process it is highly desirable to introduce as much realism as possible by use of actual data instead of subjective assumptions.

The partial data on dependence that are usually available, which can be used by the MSF model, are the number of times each combination of failures is observed without any information on total number of opportunities. That is, for a three-unit system, the partial data necessary are the number of times the human has made errors in one of the units, in two of the units, and in all three of the units. From LER data bases or other test- and maintenance-related records in nuclear power plants, this information can be gathered, whereas the information on total number of opportunities is rather difficult to obtain (see NUREG/CR-3519). The independent failure probability (p) for the

actions involved can be obtained from the present human error rate (HER) being used. The MSF model is capable of providing an estimate of the dependent failure probability using the partial data on dependence and knowledge of the independent failure probability. Figure A-2 provides the steps involved in performing the calculation using the MSF model.

3.2.1 Example Calculation

Consider the situation in a two train containment spray injection system (CSIS) where both CSIS pump flow recirculation valves are tested monthly. These valves, which are normally closed, are opened for testing. There is a possibility that owing to dependence of action following the test, both valves may be left open. The system success requirement is 1 out of 2:G logic (i.e., at least one of the valves should be in its correct normal position for successful operation of the system). For performing system unavailability calculations, the probability of leaving both valves open must be calculated. In this situation, partial data available to the analyst.

<u>Step 1 - Determine the Number of Dependent Actions Involved</u>

This step is similar to step 1 in Section 3.1. In the present example problem, the number of actions involved is two (i.e., $r = 2$).

<u>Step 2 - Determine the System Success Configuration</u>

This step is similar to step 2 in Section 3.1. In the present example problem, the system success configuration is 1 out of 2:G logic.

<u>Step 3 - Collect Available Data on Dependence</u>

In this situation, since partial data are available, these data are collected and structured in a format that can be used by the MSF model.

The data should be structured in the following manner for an r-unit system.

Number of Channels/Units Affected	Number of Times
1	x_1
2	x_2
:	:
r	x_r

where x_i denotes the number of times i out of n units were in failed condition owing to operator action.

In the present example problem, the available data for the 2-unit system are assumed to be the following: (x_1=95, x_2=11).

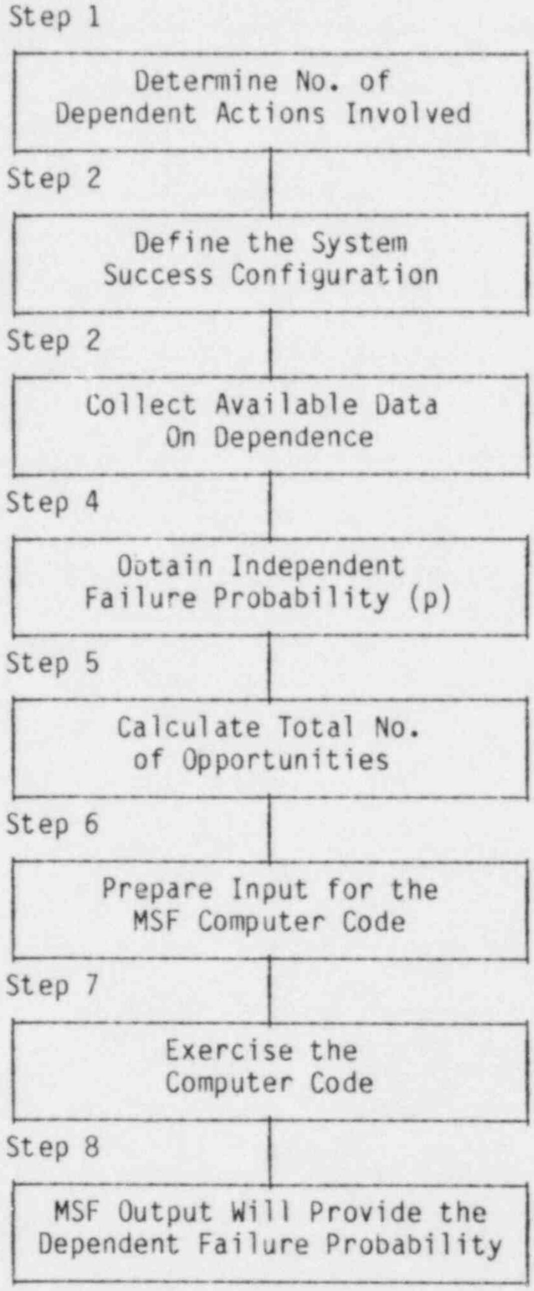


Figure A-2 Steps involved when partial data are available and p is known.

Number of Channels/Unit Affected	Number of Times
1	95
2	11

Step 4 - Obtain Independent Failure Probability (p)

This step is similar to step 3 in Section 3.1 of this appendix. In the present example problem, the independent failure probability of leaving the CSIS pump open following the test is 1×10^{-2} ($p=10^{-2}$).

Step 5 - Calculate Total Number of Opportunities (N)

On the basis of the independent failure probability (p), the partial data will be complimented by calculating the total number of opportunities. Let x_0 be the number of times the operator performs all the actions correctly. Then the total number of opportunities (N) is given by

$$N = x_0 + \sum_{i=1}^{r} x_r$$

where the second term is obtained from the previous steps and x_0 is obtained from the equation

$$(1-p)^r = \frac{x_0}{x_0 + \sum_{i=1}^{r} x_i}$$

In the present example problem,

$$r = 2,\ p = 10^{-2},\ x_1 = 95,\ x_2 = 11,$$

$$x_0 = \frac{(95 + 11)\ (0.99)^2}{[1 - (0.99)^2]} = 5270,$$

and N = 5376.

Step 6 - Prepare Input for the MSF Computer Code

The input to the MSF Computer code can now be prepared with the information gathered in the five preceding steps. Appendix C provides instructions for preparing the input for MSF code and it consists of two input cards.

Step 7 - Exercise the MSF Computer Code

The step involves running the computer code with the input deck prepared.

Step 8 - Obtain Results from the MSF Output

The computer code performs all the computations necessary for obtaining the dependent failure probability. It computes the degree of dependence (k) from the data and provides results for different system success configura-

tions. The analyst will choose the appropriate result corresponding to the system success configuration defined in step 2. For the example problem, the dependent failure probability for both valves is 2.1×10^{-3}, i.e.,

$$^{1}P_{2} = 2.1 \times 10^{-3}$$

3.3 Dependencies Between Different Units Are Similar and Complete Data on Dependent Failures are Available

This is the ideal situation for calculating the dependent failure probability when the data requirement of the model is completely satisfied and the analyst does not require any subjective assumption. At the present time, however, actual data to this level are not available. Nevertheless, structured expert judgment opinion can be introduced. If survey or interviews are conducted with the personnel responsible for carrying out the tasks, the responses will also be obtained in terms of complete dependent failure data, which in turn will be used by the MSF model to derive the dependent failure probability desired.

The complete data to be collected include the total number of opportunities and the number of times each combination of the units were affected by the operator. In this case, all the necessary computations will be conducted by the MSF computer code to provide the dependent failure probability. Figure A-3 provides the steps involved in performing the calculations in these scenarios.

Step 1 - Determine the Number of Dependent Actions Involved

This step is similar to step 1 in Sections 3.1 and 3.2 of this appendix.

Step 2 - Define System Success Configuration

This step is similar to step 2 in Sections 3.1 and 3.2 of this appendix.

Step 3 - Collect Available Data

In this step the data should be structured in a manner usable by the MSF model. This is similar to step 3 in Section 3.2 of this appendix except that additional data are required on total number of opportunities.

Number of Channels/Units Affected	Number of Times
1	x_1
2	x_2
:	:
r	x_r
Total number of opportunities	x_0

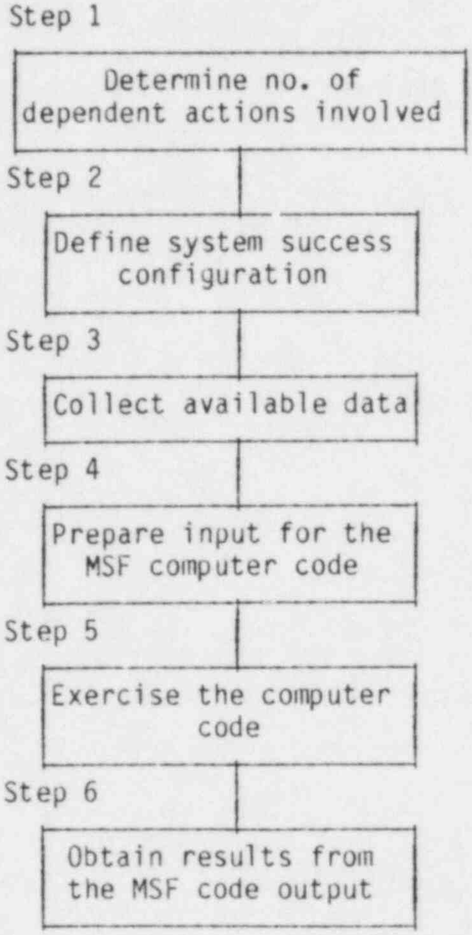


Figure A-3 Task flow diagram when complete data are available.

Data of this type can be gathered through similation or searching of actual test, maintenance, and calibration records. When actual data are not available, and structured judgment opinion, interview of personnel involved in the action, or the judgment of a reliability analyst is to be used, this would be the appropriate level. That is, individuals will be more comfortable and probably more accurate in evaluating the dependency situation in terms of this data of this type.

Step 4 - Prepare Input for MSF Computer Code

This is similar to step 6 in Section 3.2 of this appendix. Direction for preparation of input deck is provided in Appendix C. The input deck for the MSF Code consists of two records.

Step 5 - Exercise the Computer Code

The MSF Computer Code shall now be exercised with the input deck prepared. The computer code will perform all the calculations necessary and

provide the dependent failure probability for possible reliability configuration of the system being considered.

Step 6 - Obtain Results from MSF Output

The appropriate dependent failure probability for the system success configuration defined in step 2 can be obtained directly from the MSF computer output.

3.4 Dependencies Are Different Within Subset of Tasks and the Degree of Dependencies Are Directly Estimated from Structured Expert Judgment

In many situations, the dependent failure situation does not directly reveal itself as a clear-cut failure/success configuration. However, such situations can always be divided into subsets of tasks, which in return possess clear-cut failure/success configurations. If the system fault trees of which the dependent failures are part are developed in sufficient detail, the dependent failures will be clear-cut failure/success configurations. Otherwise, the user shall split the overall dependent condition into subset of tasks. Usually, the degree of dependence within the subset of tasks is expected to be different from that among the subtasks. In such situations, the analyst will have to proceed by defining the overall tasks into a subset of tasks and the calculation will proceed in steps using the MSF model as outlined in the following. This is described in a stepwise manner in Figure 3.4.

Illustration:

Consider the situation where the dependent failure probability of leaving each of three pairs of auxiliary feedwater system pump discharge valves closed following monthly tests is to be determined. The probability of leaving the valve closed following test is 10^{-2} and a strong dependence is expected for each pair of valves. Also, since the three pump tests are performed sequentially as part of the same general procedure, the three faults are coupled to some extent.

Step 1 - Define Logic Configuration of the Overall Dependent Failure

The overall dependent failure logic configuration shall be defined in terms of subset of tasks consisting of dependent failures. This can be explained through an example. It is similar to performing a very simple fault tree analysis of the dependent failure. For the example problem, the dependent failure situation is the failure to leave each of three pairs of valves open following test. The logic configuration is depicted in the diagram below:

Step 2 - Define Subtasks with Dependency Within It

Once the overall logic configuration of dependent failures is defined, the subtasks will reveal themselves. These subtasks are themselves dependent failures that need quantification for determination of the overall dependent

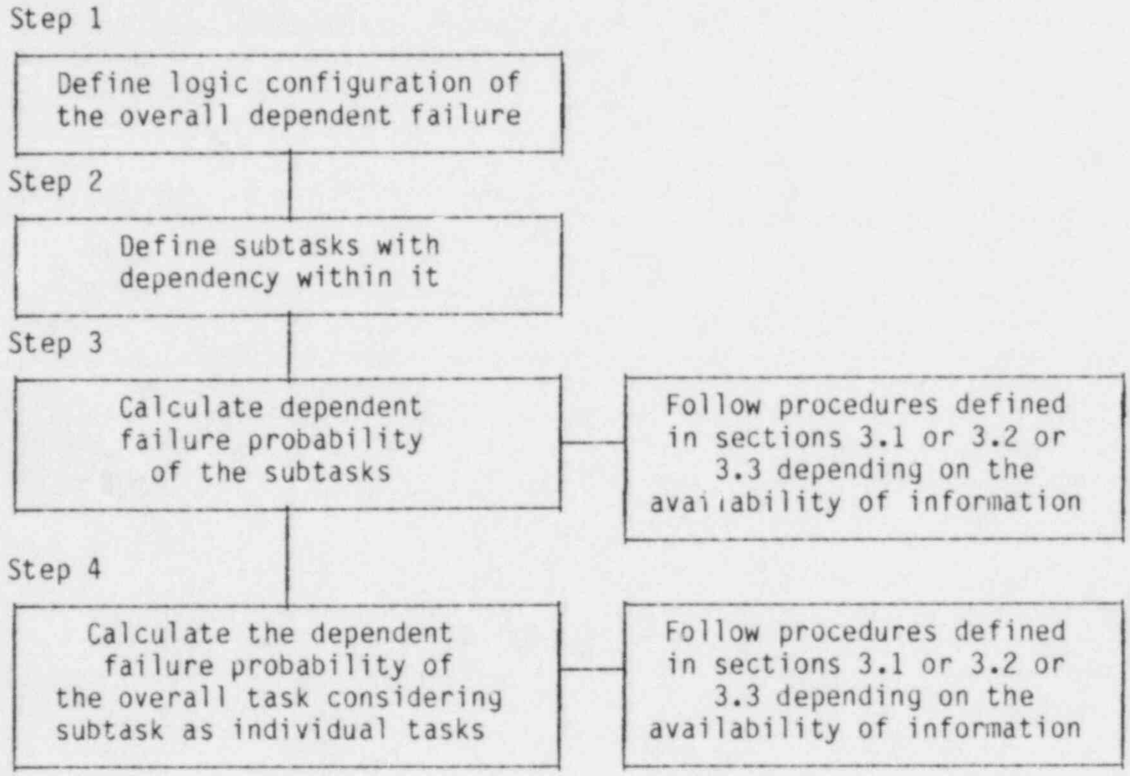
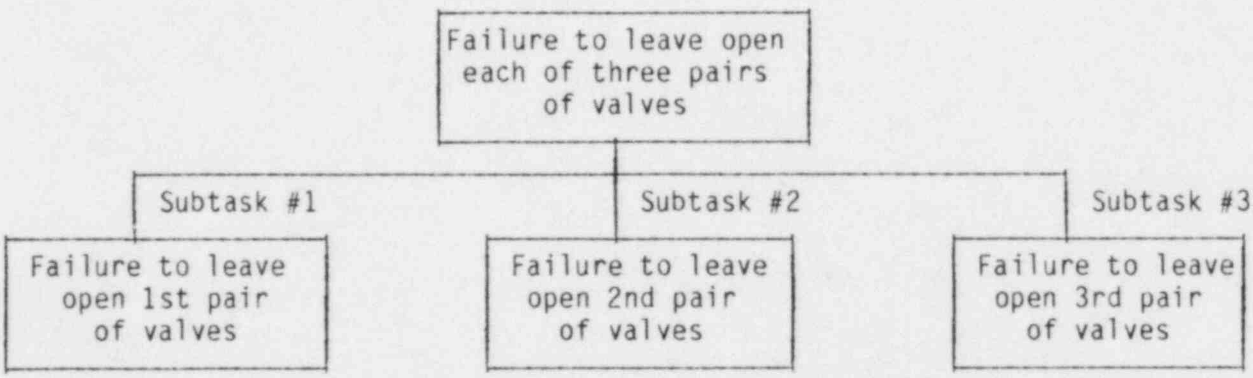


Figure A-4 Flow diagram when κ is different within subset of tasks.



failure probability. In the example problem, the subtasks are the failures to leave open a pair of valves. In this particular case all three subtasks are equal and the dependent failure probability of each will be the same.

Step 3 - Calculate the Dependent Failure Probability of Subtasks

The calculation of dependent failure probability of each of the subtasks can be performed using the MSF model following procedures defined in Section 3.1, 3.2, or 3.3 of this appendix depending on the type of data available. For this example problem, since all three subtasks are the same, the calculation needs to be performed only once. Let us also assume that no data

are available, and the analyst will perform the calculation using his knowledge of independent failure probability (p) and an estimate of the degree of dependence (k) from structured judgments. Accordingly, procedures defined in Section 3.1 of this appendix will be followed. For the subtasks $p=10^{-2}$, and let us assume that k is estimated to be 0.9. This is a 2-unit system with 1 out of 2:G logic and the dependent failure probability is

$$
\begin{aligned}
{}^{1}P_{2} &= kp + p^{2}(1-k) \\
&= 8.02 \times 10^{-3}
\end{aligned}
$$

Step 4 - Calculate the Dependent Failure Probability of the Overall Task

Once the dependent failure probability of each of the subtasks has been calculated, the overall dependent failure probability shall be calculated using the MSF model considering each of the subtasks as the basic task. This can be performed following either of the procedures described in Section 3.1, 3.2, or 3.3 of this appendix. In the example problem, the system success configuration is 1 out of 3:G logic. The independent failure probability is 8.02×10^{-3}, as calculated in the previous step, and let us assume that k is estimated to be 0.2. The dependent failure probability using the procedures in Section 3.1 of this appendix is

$$
\begin{aligned}
{}^{1}P_{3} &= \sum_{i=o}^{2} \left[1-(1-p)\ (1-k)^{i}\right] \\
&= p\ [1-(1-p)\ (1-k)]\ [1-\ (1-p)\ (1-k)^{2}]\ , \\
&= 6.1\ \times 10^{-4}\ .
\end{aligned}
$$

APPENDIX B

ERROR DEPENDENCE IN REPETITION OF TASKS WITHOUT FEEDBACK[1]

H. William Morrison[2]
Department of Psychology
State University of New York at Stony Brook

1.0 INTRODUCTION

Applied studies of performance errors have apparently focussed more on gross error rates or error/productivity tradeoffs than on dependence among errors. However, dependence among human errors could increase the probability of simultaneous failure in redundant subsystems intended to produce a practically fail-safe system (NUREG/CR-2211, 1981; NUREG/CR-1278, 1983). To the extent that errors in human performance contribute to dependent failures in parallel components or channels, redundancy's purpose is defeated. Dependence may not be of much practical concern if errors are immediately detected and corrected, are routinely handled through quality control or quality assurance procedures, are almost certainly discovered in time to forestall serious consequences, allow recovery after halting a malfunctioning process, and/or produce failures which can be tolerated. However, human errors in test, calibration, and maintenance tasks may not generate immediate feedback, may not be readily subject to inspection or testing, and may only be detectable by repeating the entire procedure (with its own risk of error). System failures which resulted from such dependent errors might be difficult to tolerate or terminate gracefully in the operation of nuclear power plants, processing of hazardous substances, space missions, and military systems.

The psychological literature suggests that feedback (or knowledge of results, or reinforcement) is crucial in shaping an individual's performance, and without feedback any error might be perpetuated indefinitely, representing total dependence (except for some form of random fluctuation). However, just as errors can suddenly appear, it is possible to imagine either apparently spontaneous recovery, or recovery attributable to external events. Failures of specific perceptual-cognitive-motor processes or "programs" and eventual recovery from failure might account for dependence of errors. Alternatively, variation in general psychophysiological state could produce variation in rate of (independent) errors, producing statistical dependence, or variation in environmental conditions affecting performance could produce the same result. This view disagrees with Swain & Guttmann's discussion (NUREG/CR-1278, 1983,

---

1. This research was carried out under Contract No. 122691-S with Brookhaven National Laboratory/Associated Universities, Inc.
2. Thomas Maloney and Evan Chua-Yap assisted in data collection. I appreciate the assistance of John O'Brien, Contract Liaison Officer, Robert E. Hall, and Xenia Coulter in various phases of the study. I am grateful to P. K. Samanta for computations using his moment estimation procedure and to John Richters and William Persons for analysis of variance computations. Richard Reeder's technical advice and support were invaluable.

p. 10-3) of indirect dependence. They do not consider changes in error rates of independent events as a source of dependence. But both research and application involve domains extending over time and changing values of other variables, and even if such domains could be partitioned into locally independent regions, events may be statistically dependent over the domain as a whole.[3] Thus, consideration of dependence requires care in defining the domain and selecting the level of analysis appropriate for a given application. The present study attempts to address the question of dependence of errors in an individual's performance of a task, without attempting to distinguish between possible sources of error dependence. Finally, errors could be negatively dependent with the probability of success increased following error, possibly even without recognition of an error's occurrence.

Even if all other performance factors were held constant, an inescapable characteristic of the repeated performances necessary for occurrence of dependent errors is the temporal sequence and spacing of trials. Classical concepts to account for changes in performance over time in addition to learning include fatigue or inhibition, where the probability of failure increases with continuous practice even without previous errors, and processes associated with the beginning and anticipated end of a series of trials, such as "warm up" effects or an "end spurt." Such effects of scheduling must be considered as another possible explanation for systematic change in errors, in addition to their dependence.

1.1 Models

Most if not all mathematical models of learning in an extensive research literature include conditional probability terms representing dependence of errors. These models have generally been applied to performance during the learning process, for simple tasks with feedback, and where error rates are not low. By contrast the present study is concerned with asymptotic performance, relatively low error rates, complex tasks without explicit feedback but

---

3. For example, one local region could have $p(a) = p(b) = x$ with probability of joint occurence $p(ab) = x^2$, while in a different region $p(A) = p(B) = cx$ $(c \neq 1)$ with probability of joint occurence $p(AB) = c^2x^2$, so that within each region the events are independent. In sampling joint events equally often from within each region we would have $p(a \vee A) = p(b \vee B) = (x+cx)/2$ and $p(ab \vee AB) = (x^2+c^2x^2)/2$, producing dependence because the latter is not the same as $p(a \vee A)p(b \vee B) = (x+cx)^2/4$. This problem of heterogeneous domains recurs under various labels in statistical analysis, but requires attention here because decisions must be made about combining data from distinguishable subdomains, such as different people and/or tasks, which could produce apparent dependence between errors even if within each sub-domain they were independent. The present study, concerned with dependence in an individual's performances of a task, nevertheless analyzed data pooled over individuals and/or tasks because of statistical considerations (and with a recognized risk of finding spurious dependence). From the point of view of another domain, such pooled data would be entirely appropriate for assessing dependence at that level of analysis.

where recognized errors can be corrected, and under conditions simulating some features of actual work places.

The Multiple Sequential Failure model proposed by Samanta & Mitra (NUREG/CR-2211, 1981) for error dependence (including performance of test, calibration, and maintenance tasks) is a simple (n+1)-state Markov process. The state i (i=0,1,...,n) represents the number of successive previous errors. The model has two parameters, the base error rate p (termed the "independent failure probability" by Samanta & Mitra, and the "basic human error probability" in NUREG/CR-1278, 1983) and the dependence factor k (with the values of both parameters in the interval zero to one). For the case where all tasks are the same the model proposes that the probability of the i-th error and transition to state i, given that i-1 successive errors have just occurred, is

$$1) \qquad p_i = \begin{cases} p & \text{if the current state } i-1 = 0 \\ p_{i-1} + k(1 - p_{i-1}) & \text{if the current state } i-1 \neq 0 \end{cases}$$

while the probability of no error and transition to state 0 is $(1 - p_i)$. The various transition (and error) probabilities are obtained by repeated application of Equation 1. The observed gross error rate P (total number of errors divided by number of trials) depends on all these transition probabilities and in particular is not an estimate of the base error rate p unless k=0. Noteworthy qualitative features of Samanta & Mitra's model are: independence of path with state determined only by the number of immediately preceding successive errors; provision for positive dependence only; a negatively accelerated increase in error rate with successive errors; and an increase in error rate following error(s) inversely related to p, but as a practical matter with overwhelming determination of repeated errors by the parameter k when values of p are close to zero. This model has the flavor of a learning model where error increasingly occurs with practice.

Alternatively, the increase in error rate with successive errors might be (initially) linear, with

$$2) \qquad p_i = p_{i-1} + \min(k\ ,\ 1-p_{i-1}) \qquad \text{if the current state } i-1 \neq 0$$

or (initially) positively accelerated, e.g., with

$$3) \qquad p_i = p_{i-1} + \min(kp_{i-1}\ ,\ 1-p_{i-1}) \qquad \text{if the current state } i-1 \neq 0$$

or sigmoidal. Although the same symbol k has been used to represent a dependence parameter in the different models, the role and value of k depends on the model, and in Equation 3 the parameter k could have values greater than one. Also, other functions could be chosen for Equations 1 and 3 to produce the same qualitative results, i.e., negatively or positively accelerated rates of increase with successive errors.

By comparison the model represented by Swain & Guttmann's (NUREG/CR-1278, 1983) procedure, leaving aside their pragmatic choice of values for k, is a two-state Markov model. Denoting the number of immediately preceding

successive errors on identical tasks by i-1, as above, the probability of the i-th successive error can be written

$$4)\qquad p_i = \begin{cases} p & \text{if the current state } i\text{-}1 = 0 \\ p + k(1 - p) & \text{if the current state } i\text{-}1 \neq 0 \end{cases}$$

while the probability of no error and transition to state 0 is again $(1 - p_i)$. In this model $p_2 = p_3 = \ldots = p_n$ and the indistinguishable non-zero states can be collapsed, producing a two-state model. Noteworthy features of this model are: stronger independence of path than Samanta & Mitra's model, with state determined only by presence or absence of an immediately preceding error; provision for positive dependence only; a constant rate of successive errors beginning with the second error; and an increase in error rate following the first error inversely related to p, but again with overwhelming determination of repeated errors by the parameter k when values of p are close to zero. The arbitrary choice of 0, .05, .143, .50, and 1 as values of k for applicaton of this model is separable from the question of the <u>form</u> of dependence.

All these models permit independence with k=0, which could be rejected if $p_1 \neq p_2$. But for a given task <u>any</u> of the models above could fit <u>any</u> estimated values of p_1 and p_2 (with $p_1 \leq p_2$) equally well. Only fitting the values of p_3 or later p_is would challenge a given model or discriminate between alternative models. Conversely, the models' implications for multiple failures differ only when more than two successive errors are involved. In extrapolations from given values of p_1 and p_2, the models' predicted probabilities of subsequent successive errors, p_3, p_4, etc., will be ordered from lowest to highest as follows: two-state model (Equation 4), negatively accelerated model (e.g., Equation 1), linear model (Equation 2), and positively accelerated model (e.g., Equation 3), unless the p_is reach 1.0 or some other limiting value. For given values of p_i on non-successive trials, interpolated p_is predicted by the various models would have this ordering reversed.

Such models might be fruitfully applied because, if valid, they simplify the problem of predicting rate of multiple performance errors by requiring estimates of only two parameters. However, just as the base error rate p will certainly vary over individuals and tasks, the dependence represented by k might also vary and have to be estimated separately for different individuals, tasks, and/or situations. Alternatively, the form of dependence might be described better by some different model, and the form of dependence itself could vary from situation to situation. But even a model demonstrably wrong in some details might serve a pragmatic or heuristic purpose in improving estimates of error dependence, leading to more accurate assessment of risk.[4]

---

4. The practical consequences of assuming the wrong model would not only depend on system design and level of dependence, but also on any systematic errors of estimation which resulted from fitting data to the wrong model, as well as on the validity and reliability of data used to estimate k. Sensitivity analyses to determine the effect of varying assumptions on calculated risk of simultaneous failure (NUREG/CR-2211, 1981; NUREG/CR-1278, 1983) could

Even though such models might apply to inter- as well as intra-task dependence (NUREG/CR-2211, 1981; NUREG/CR-1278, 1983), here we consider only dependence within repeated performance of the same task.

1.2 Empirical Research

The purpose of the present study was to collect systematic human performance data to evaluate Samanta & Mitra's model in Equation 1, and, assuming the validity of the model, to examine variation in k over individuals and tasks. Previous studies of this model (NUREG/CR-2211, 1981; Schurman & Hawley, 1982) only estimated its parameters without attempting to test goodness-of-fit. Both studies pooled results over individuals, and the second pooled results over tasks as well. As Samanta & Mitra recognized, their data from Licensee Event Reports represented self-detected and corrected errors and were subject to bias, one extreme observation was excluded from the analysis, and the analysis depended on a derived estimate of the frequency of errorless sequences assuming that p=.01. Their estimates for k ranged from .199 to .403, depending on whether they considered two, three, or four trials (all based on the same data set).

Only Schurman & Hawley's results in the upper half of their Table 1 (1982, p. 13), for Army motor pool mechanics' uncorrected errors in repetitions of the same task are pertinent here, since their other analyses concerned dependence between different stages of tasks, and self-detected and corrected errors. The authors pointed out that inclusion of errors from "trial-and-error sequences" is debatable, and their discussion suggests that the "repeated" tasks sometimes actually represented different problems (pp. 8-9). Apparently without the "trial-and-error" errors there would have been too few multiple errors to analyze. They estimated parameters for tasks performed twice (p=.333 and k=.667) and for tasks performed four times (without finding a solution, but assigning a default value of k=1.0). For both double and quadruple trials the frequency distributions of error counts were U-shaped, which could be attributed to the heterogeneity of individuals and/or tasks (with varying values of p and/or k) included in the analysis.

1.3 Levels of Analysis and Types of Errors

Behavioral accomplishment can be considered at any level from microscopic intra-organismic events to the molecular, such as pushing a button, to the molar, such as executing a complex sequence of actions. Different forms of dependence might apply to different levels of behavior in tasks, or as Swain & Guttmann point out (NUREG/CR-1278, 1983) to different kinds of errors in their execution. The level was selected here to correspond to the intended application, and performance was scored either as completely correct or

---

(cont.)
be extended and refined to take into account possible alternate forms of dependence, errors in estimating k from fitting the wrong model, and unreliability of the data upon which estimates are based.

containing error(s) without attempting to classify or relate various types of subordinate errors.

1.4 Present Approach

Presentation of actual physical tasks and recording of performance errors made by trained technicians were beyond the scope of the present study. The study utilized schematic task layouts presented by video display, with responses using a joystick and trigger which controlled a cursor, intended to resemble tasks performed by remote control with a video monitor. Three experimental tasks were designed as analogues of testing, maintenance, and/or calibration tasks in nuclear power plants, and were accepted for this purpose by Brookhaven National Laboratory.

Work characteristics to be simulated in the experiment were: routine, repetitive performance; intermittent scheduling interspersed with other tasks; well-practiced subjects with relatively low error rates (intended to be less than 10%); no direct feedback on errors and no direct supervision of work; continuous availability of written procedures for each task; and the option of restarting a task without penalty.[5]

Each task was presented in blocks of three successive trials, both to maximize presumed dependence, and to collect as much data as possible in each session. In addition, an "artificial" Composite Task (of which subjects were unaware) was defined by joint outcomes of the three different tasks within a three-task cycle, with error on the composite task defined as an error on any of the three different tasks in corresponding positions within their blocks.

2.0 METHOD

Five subjects repeatedly performed three tasks in sessions usually lasting about four hours, over a period of one month.

2.1 Subjects

Eight male Stony Brook students were recruited by a campus newspaper advertisement. Scheduling problems severely constrained selection of subjects. One subject was subsequently dismissed because of high error rates, and two subjects had to withdraw (because of an automobile accident and

---

5. Recognized differences from actual tasks include: shorter durations of interpolated tasks and a smaller repertoire of tasks; no check-off of task steps, except for signifying task completion; differences in motivation both in terms of the "demand character" of an experiment, and absence of potential serious consequences of errors; performance under conditions not considered hazardous; work perspective of a temporary rather than a permanent employee; possibly more concentrated activity without interruption or change in physical location, and without pay for "breaks;" and the recognition that performance records were being kept. Furthermore student subjects and technicians represent populations different in many obvious respects.

because of eyestrain). The five remaining undergraduate students were 19, 22, 23, 26, and 33 years of age; were majoring in Physics, Psychology (3), and Religious Studies; reported cumulative Stony Brook GPAs of 3.1-3.7 (for the three reporting), SAT verbal scores of 490-670 (for the four reporting), and SAT quantitative scores of 610-760 (for the four reporting).

2.2 Tasks

Three analogues of test, calibration, and/or maintenance tasks were presented by an IBM PC computer with color graphics board and Zenith or NEC monochromatic (green on black) 12" monitor; responses utilized a Hayes Mach III joystick with button trigger. Schematic task displays were presented in character mode on the 25x40 character screen. Each subject adjusted screen intensity to suit himself. The Advanced Basic interactive program which controlled the administration of tasks was executed interpretively, producing a sluggish cursor response to joystick actions. Correct performance required a prescribed sequence of actions or responses (some contingent on varying displayed conditions), using the joystick to control movement of the blinking cursor to response locations represented by reverse image figures, and pressing the joystick trigger once to "highlight" the cursor location until the trigger was released. Where different functions were prescribed for the same location (e.g., opening or closing a valve), different responses were not required (or permitted). A lit indicator was represented in the display by a "highlighted" reverse image symbol.

The Attachment presents procedures, photographs of displays, and supplementary descriptions for the three tasks: Task 1, "Valve Isolation and Check," required at least 30 actions for correct completion (not counting "end" which signified task completion); Task 2, "Wiring and Metering," required at least 38 actions; and Task 3, "Logic Sequence Tests," required at least 32 actions. Free options within correct performance of a task were: delay in responding, duration of trigger press (restricted in some cases), restarting a task from the beginning, pausing to review the instructions and subsequently resuming or restarting the task, moving the cursor without pressing the trigger, or pressing the trigger when the cursor was on a blank screen position (restricted in Task 2). Pressing the trigger with the cursor marking a non-response figure was counted an error, "damaging equipment."

During data collection any other actions were accepted and responded to by the program, and counted an error, without any indication of error to the subject. Possible exceptions where implicit feedback may have been recognized were: in Task 1 if the subject ended the task (instead of opening or closing the cover) after completing the first or second phase of the task; in Task 2 if the wiring area (including the terminals to which connections were made) was cleared prematurely; in Task 1 or 2 if a meter or indicator required action and the subject realized he had not been monitoring its value; and in Task 3 if the subject (mistakenly) planned to respond in a block where neither indicator was lit. Only the first of these did not offer the remedy of restarting the task if necessary to correct the error.

2.3 Instructions and Setting

Subjects were told that the study concerned repeated performance of tasks, that reducing or eliminating errors was of primary importance, but that they should increase productivity to the extent possible without increasing errors. The task procedures (included in the Attachment) were always available and subjects were told that they could use or refer to them at any time, but not to make or work from other notes. Apparently subjects seldom referred to the procedures during data collection. They were encouraged to restart a task (without its counting as an error and with no effect on their pay or performance record except for "lowered productivity") if they had made an error or were confused. Participation was described as a part-time temporary job, with pay increasing from $3.35/hr. by approximately $.20/hr. with each succeeding session.

To the extent possible two subjects were scheduled together for each four hour period, and worked in the same room at desks separated by a partial partition. An experimenter began each session, but usually left the room during the session although he was "on call" in case of difficulties. Subjects could cancel or reschedule sessions by prior arrangement, and could terminate a session early, but were not permitted to schedule more than one session per day.

Subjects were told they could converse as long as it did not interfere with their work, but were asked not to discuss the tasks and not to eat, drink, smoke, or use audio equipment in the room. Before another trial on the same task subjects could review the procedure before continuing (as they could during a task), but during the inter-task interval subjects sometimes paused to "stretch" whether or not they indicated that they were reviewing. Between presentations of different tasks subjects could take a "break" (the only time they were supposed to leave the room, and the only time for which they were not paid) or terminate the session, but occasional periods of inactivity without indicating a break were recorded.

Although an experimenter was not present to monitor their work, it was obvious to subjects that the computer could and did record performance data. Those subjects who completed the experiment appeared to be conscientious in their work. They seemed to consider the experiment and working conditions a temporary part-time job. They reported informally that serving in the experiment was the most boring work they ever did, with several complaining that the tasks did not permit them to move around and several complaints about having to attend continuously during the tasks, as well as a complaint about eyestrain from one subject who completed the experiment.

2.4 Training and Feedback

During initial training the computer did not act on an erroneous reponse except to highlight the cursor position and to provide a feedback message at the bottom of the screen; the subject could not proceed before he performed the correct action (except for feedback on errors in duration of trigger

press, which subjects were not required to correct or repeat).[6] When error rates failed to decrease satisfactorily during training, subjects were told to "shape up on errors" and "also productivity, to the extent possible." The final stage of training omitted feedback during a trial, and only after completing each task were subjects informed what their first error had been. After an initial orientation subjects completed eight or nine training sessions prior to data collection, except for one late recruit who had six sessions. During data collection subjects received no information on performance errors.

2.5 Schedule

A block of three repetitions of a task was presented before the next task was randomly selected, subject to the restrictions that the same task was not presented in successive blocks (within or between sessions) and that within successive three-block cycles each task was presented once.

If a subject had not already chosen to end the session, the control program terminated the session when there was apparently not enough time remaining to complete another block within the allotted time. Useable data from eight to eleven sessions were collected for each subject. Data from two subject-sessions were discarded because of incorrect calibration of the joystick.

2.6 Scoring

Performance of a task was scored correct only if all prescribed actions were performed with no errors of omission, commission, sequence, or response duration. The Composite Task was scored correct only if (within a cycle of three blocks) the three different tasks in corresponding block positions were all performed correctly.

3.0 RESULTS

In addition to formal quantitative analyses, enough details are included to provide some sense of the data, which should help avoid a too abstract view of results. Particularly in attempts to test goodness-of-fit, confidence in the results would have benefitted from more observations.

3.1 Stability Over Sessions

Table B.1 reports the gross error rate and mean time on task for all completed tasks (for final performance on each trial, ignoring review time and performance cancelled by a restart) for successive sessions, for each subject and task during data collection when feedback was omitted. These results describe general subject and task performance characteristics (including any

---

6. This training also provided a partial check of the control program's agreement with the instructions, although the entire program was not exercised in this phase.

effects of dependence), and can be examined for trends which would contradict the assumption of stable performance over the period of data collection. Consistent trends over sessions do not appear in either error rate or time. The high gross error rates for Subject 5 suggest that his results should be treated separately in analyses or data pooled over subjects.

3.2 Stability Within Sessions

Error probabilities might have depended on whether a block occurred first, midway, or last in a session, and/or on a trial's position as first, second, or third in a block, possibly confounding subsequent estimates of dependence. The error rate was calculated for trials in each of these positions, omitting trials from the calculation if an error occurred on the previous presentation of the same task in that session.[7] According to the models' independence of path assumptions, omitting these trials excludes effects of dependence from the results. Substantial differences in estimates of p appeared for first, second, and third trials within a block, with (unweighted) means (over subjects, tasks, and block's position in session) of .087, .097, and .164, respectively, suggesting a "third trial" or "end effect" within blocks, possibly from fatigue or motivational factors.

Repeated measures analyses of variance of the unweighted proportions were performed on the factorial layout of 3 tasks x 3 block positions in session x 3 trial positions in block, both for estimates of p and log (p + .01), for Subjects 1-5 and for Subjects 1-4. In all four analyses the differences between tasks were significant (in one case near significant after adjustment for departure from statistical assumptions). For the logged data of all five subjects the differences between trial positions were significant with $F(2,8)=4.97$, and marginally significant around the .10 level in the other three analyses. Of all the other tests only two, of different interactions in different analyses, approached significance (.10).

Whether or not the trial position differences were significant, the presence of such position effects could, compared with the presumed effects of dependence alone, inflate the frequencies of the within-block error sequences 001 (i.e., correct-correct-error, one of the three single error sequences), 011 and 101 (two of the three double error sequences), and 111 (the only triple error sequence) and thus inflate within-block estimates of p_2 relative to p_1, and particularly p_3 relative to p_1 and p_2. The (weighted) proportion of errors in each block position, for each subject and task including the Composite Task, are reported in Table B.6 in the Attachment, for possible use in interpreting subsequent estimates of parameters.

3.3 Direct Estimation of p_is

The clearest overview of dependence is provided by direct estimates of the conditional probabilities of error, $p(1/0)=p_1$, $p(1/01)=p_2$, $p(1/011)=p_3$,

---

7. Seven trials from incomplete blocks at the ends of sessions were also omitted from these and subsequent analyses.

Table B.1 Stability Over Sessions of Proportion of Errors (P) and Mean Time on Task in Minutes (T) For Final Task Performance

Subject	Task		Session 1	2	3	4	5	6	7	8	9	10	11
1	1	P	0.0	0.0	0.0	.11	0.0	0.0	0.0	.11	.11		
		T	3.84	3.92	3.47	3.65	3.60	3.71	3.75	3.54	4.02		
	2	P	0.0	.11	0.0	0.0	0.0	0.0	0.0	0.0	.11		
		T	6.19	5.99	5.23	5.70	5.97	5.37	5.61	5.01	5.77		
	3	P	.11	0.0	.10	.33	0.0	0.0	0.0	0.0	0.0		
		T	6.84	7.08	6.65	7.13	6.76	6.53	6.68	7.00	7.20		
2	1	P	0.0	0.0	0.0	.11	0.0	0.0	0.0	0.0			
		T	3.41	3.36	3.33	3.64	3.32	3.17	3.11	2.91			
	2	P	.11	0.0	.17	.11	.10	0.0	.11	0.0			
		T	5.02	4.83	5.55	4.94	4.31	4.60	4.71	4.36			
	3	P	0.0	.22	0.0	.22	.33	.17	0.0	.33			
		T	6.18	6.25	6.41	6.70	6.18	6.11	5.90	6.25			
3	1	P	.11	0.0	0.0	0.0	0.0	0.0	0.0	0.0	.11	.22	0.0
		T	4.21	4.02	4.59	4.15	3.98	3.80	4.27	3.39	3.82	3.86	3.97
	2	P	0.0	0.0	0.0	0.0	0.0	.11	.11	0.0	.11	.22	0.0
		T	6.28	6.10	7.46	5.91	5.49	6.08	6.58	6.42	6.45	7.08	7.09
	3	P	.11	.11	0.0	0.0	.11	.11	.17	.11	.11	.11	.11
		T	6.94	6.18	7.53	6.26	6.05	6.60	6.99	7.04	7.15	6.85	7.12
4	1	P	0.0	0.0	0.0	0.0	0.0	.33	0.0	.33	0.0	0.0	
		T	3.90	3.59	3.54	4.42	4.03	4.22	3.55	4.00	4.21	3.99	
	2	P	1.00	.33	0.0	.17	0.0	.22	0.0	0.0	.22	0.0	
		T	7.72	9.67	7.39	7.78	6.91	7.36	7.35	6.05	5.64	7.37	
	3	P	.33	0.0	.08	.17	.11	.56	.33	.33	.17	1.00	
		T	7.84	8.72	7.54	8.67	8.62	8.64	6.88	9.33	7.05	5.74	
5	1	P	.11	.22	.17	.22	0.0	.50	.33	.33	.22		
		T	4.19	4.20	4.12	3.54	3.51	3.75	4.72	4.49	4.05		
	2	P	.11	.33	.67	.27	.33	.33	.22	.89	.33		
		T	6.66	5.48	8.12	5.53	5.40	6.01	6.32	4.97	4.74		
	3	P	.33	.27	.67	.44	.30	.33	.44	.17	.22		
		T	6.56	6.88	7.27	6.80	5.73	7.18	6.80	8.15	6.85		

etc., from the entire sequence of responses within a session for each task, temporarily setting aside concerns about stability of error rates within blocks, uncertain dependence between blocks, and dependence between overlapping sequences. These analyses are based on the trial outcomes reported in the attached Table B.7 (which permits alternative analyses or computation of supplementary statistics for the error sequences). For this analysis it was again assumed that before a session's first presentation of each task the subject was in the models' state 0,[8] and following correct performance was in state 0.

All subjects had positive values of p_1 for the three individual tasks and the Composite Task. However, because no successive errors occurred in their data, the estimated value of p_2 was zero for Subjects 1-4 on Task 1, Subjects 1-3 on Task 2, and Subject 3 on Task 3, and p_3 could not be estimated since its condition was never met. If these results suggest anything, considering the small number of opportunities for successive errors with base error rates of .015-.094, then they might suggest negative dependence.[9] Furthermore, there were no occurrences of three successive errors in the data for Subjects 1-5 on Task 1, Subjects 1-3 on Tasks 2 and 3, or for Subject 3 on the Composite Task.

Estimated values of conditional probabilities for the cases with $p_2 > 0$ are shown in Figure B.1. For Subject 3 on the Composite Task (denoted S3-CT) and S5-T3 and S5-CT, p_2 is no greater than p_1, contrary to expected effects of positive dependence. However, the relative estimates of p_1 and p_2 do suggest dependence for the remaining nine cases, as do estimates of subsequent probabilities for S5-CT.[10]

In six of these cases (with p_1 estimates of .054-.192) the estimate of p_3 is based on only one or two observations, too few to compare relative fit of the alternate models (for S1-T3, S1-CT, S2-T3, S2-CT, S4-T2, and S5-T1).

---

8. Except for Subject 5 in no case was an error on the last performance of a task followed by an error on its first presentation in the next session.
9. Denoting the number of errors E and the number of correct performances C, the probability of no successive errors assuming independence can be expressed in terms of permutations as ${}_{C}P_{E-1}/{}_{C+E}P_{E-1}$. This probability was .441 for Subject 3 on Task 3 and .683 for Subject 4 on Task 1, and above .8 for the remaining subject-task combinations with no successive errors.
10. Each model predicts a specified linear regression of p_2 on p_1 for a given value of k. These lines fall on the diagonal of the plot of p_1 and p_2 when k=0, and otherwise above the diagonal, with the line passing through (0,0) for Equation 3, through (1,1) for Equations 1 and 4, and parallel to the diagonal for Equation 2. In the scatterplot of estimated values of p_1 and p_2, nine of the ten p_1 values below .10 have p_2=0, below the diagonal, balanced by the points with p_1 greater than .10 which mostly fall above the diagonal. Instead of attempting to test the hypothesis of k=0 for the entire set of observations, probably based on the assumption of a common regression line for all cases, such tests with fitted individual and task parameters are reported in section 3.8 below.

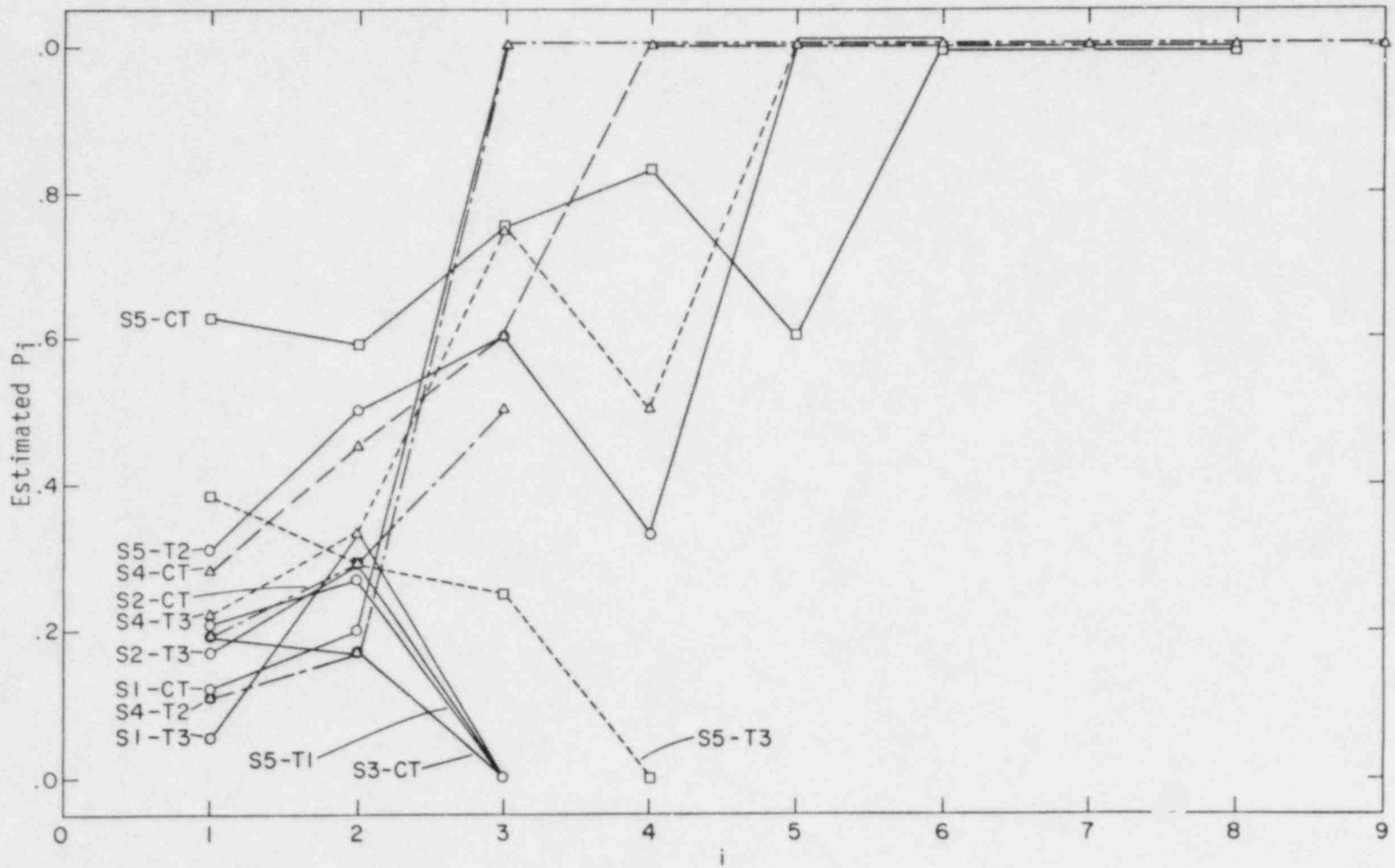


Figure B.1 Estimated conditional probability P_i following i-1 errors (all data).

For the three remaining cases (with estimated p_1 values of .219-.315) with estimates of p_3 based on four or five observations, the patterns of p_1, p_2, and p_3 estimates suggest agreement with a negatively accelerated model (e.g., Samanta & Mitra's model in Equation 1) for S5-T2, the linear model (Equation 2) for S4-CT, and a positively accelerated model (e.g., Equation 3) for S4-T3.[11]

The Figure also reveals long runs of errors for S4-T2, S4-T3 (and as a direct result for S4-CT), S5-T2, and S5-CT. Without formally evaluating their probability, these sequences seem unlikely departures from independence under any reasonable assumptions about base error rates. The fluctuations in higher p_is and their return to zero should not be counted as evidence against dependence. In the present experiment, the longer error sequences did come to an end, or were interrupted by the end of a session, but except for S5-CT the estimate of each higher p_i (above p_3) is based on only one to three observations.

A disproportionate number of breaks in error sequences occurred at the beginning of a new block, suggesting that interpolated tasks reduced dependence, and Figure B.2 presents the same p_1, p_2, and p_3 estimates omitting observations for a block's first trial if there was an error on that task's previous presentation in the session.[12] Where p_2's earlier estimate was greater than zero, the estimated values of p_2 in Figure B.2 are increased--except for S4-T3, and all these cases now have estimates of p_2 greater than p_1--except for S5-T3. For what it is worth, in those cases where p_3's estimate is based on three to five observations in this analysis, the patterns of p_1, p_2, and p_3 estimates in Figure B.2 now agree most closely with a negatively accelerated model (e.g., Samanta & Mitra's model in Equation 1) for S4-CT, a positively accelerated model (e.g., Equation 3) for S4-T3 and S5-CT, the two-stage model (Equation 4) for S5-T2, and independence for S5-T3. Actually, the changes in classification of S4-CT and S5-T2 (compared with Figure B.1) are in the same direction, toward a more negatively accelerated increase in p_i.

The apparent absence of dependence in about half the cases, frequently without any successive errors, together with long successions of errors for four subject-task combinations, is striking. While the absence of successive errors might be attributed to low base error rates of .015-.094 together with

---

11. Unlike the results in subsequent within-block analyses, these sequences of triple errors used in estimating p_3 are not necessarily confounded with any end or third trial effect, as they are when the third error necessarily occurs on the third trial within a block. Nevertheless, on the separate tasks 75% of the first three successive errors did occur within a block for both Subjects 4 and 5, while on the Composite Task 80% fell within a block for Subjects 1-4 but only 33% for Subject 5.
12. However, observations after such omitted trials were still counted as before. Factors other than reduced dependence might have contributed to error runs ending between blocks, e.g., an increased chance of error on the third trial which would reduce the chance of ending a run between the second and third trials.

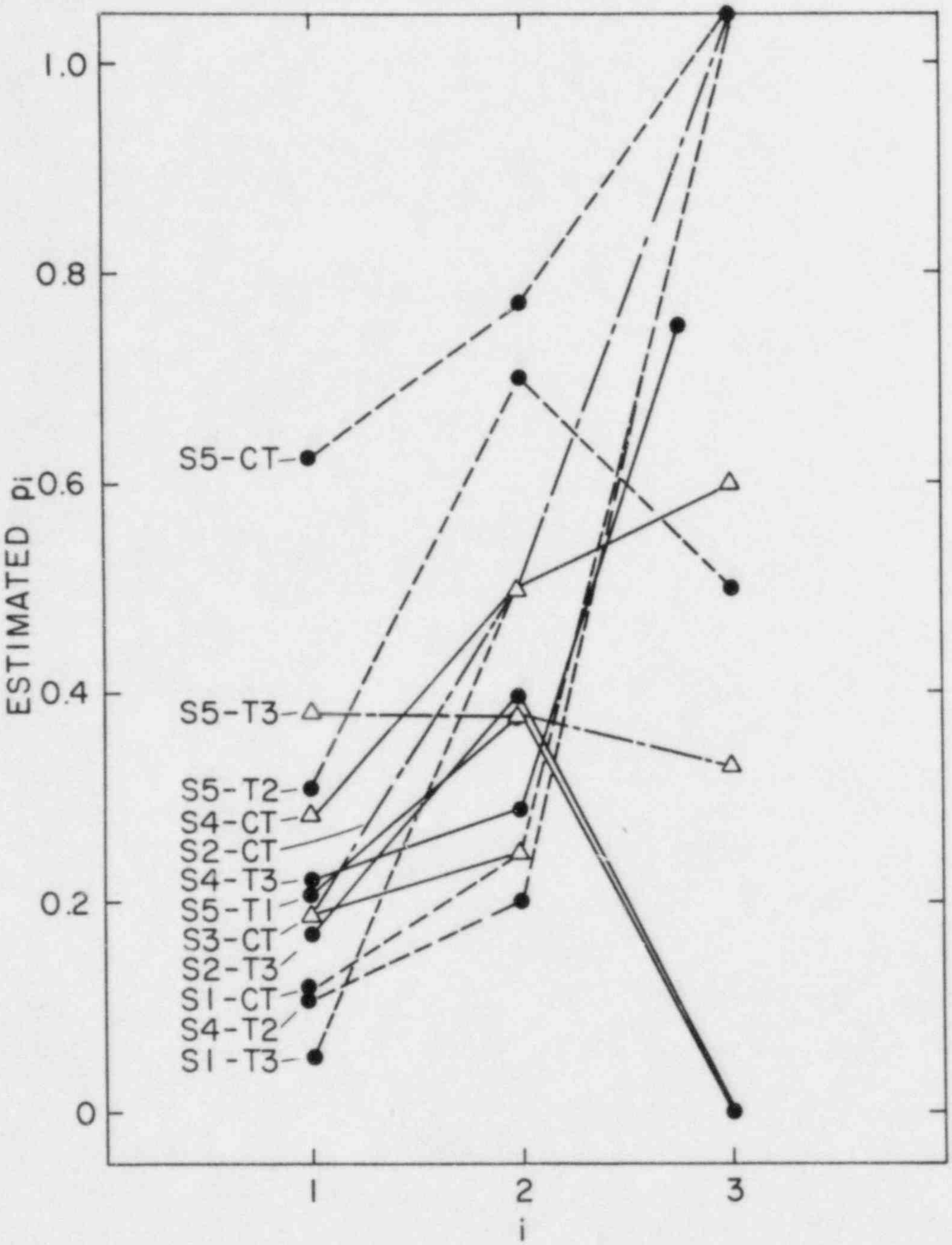


Figure B.2 Estimated conditional probability P_i following i-1 errors (excluding data for Block's first trial when an error occurred in the preceding trial).

absence of strong positive dependence, it is possible that dependence was suppressed by conditions in the experiment. Except for Subject 5 with estimated base error rates of .210-.377 on the individual tasks, the pattern of conditional probabilities suggests dependence only where either Subject 4, Task 3, or the Composite Task was involved, with estimated values of p_1 between .054 and .219 for the individual tasks and .118-.283 on the Composite Task. Where the estimate of p_3 was based on enough observations even to consider comparing relative fit of the various models, the results either suggest that different models fit different cases, or could be attributed to instability resulting from the small numbers of observations.

Subsequent analyses tested departures from independence and estimated parameter values, weighting the observations appropriately in terms of frequencies. However, both the sensitivity of these goodness-of-fit tests and the correspondence between their nominal and actual significance levels, as well as the stability of parameter estimates, could be questioned because of the small observed frequencies of multiple errors within blocks.

3.4 Tests of Independence and Parameter Estimates Within All Blocks

The results in Table B.2 and subsequent analyses follow a standard pattern described here. In order to justify the statistical assumption of independent, non-overlapping observations they are based on the eight possible within-block error sequences (000, 001,...,111). Tests of departures from independence (with k=0) in Table B.2 could be inflated by position effects within blocks, but not by between-block dependence if the tested hypothesis were true. However, the estimates of parameters on Lines 2 and 6 and tests of goodness-of-fit for the Equation 1 model reported on Lines 7-8 in Table B.2 may be distorted by dependence between blocks, a problem remedied in subsequent analyses of edited data. In any case the table allows comparisons of estimates of p and k (on Lines 2 and 6) obtained by several methods described below.

For each subject-task combination in Table B.2 the observations are frequencies of the eight possible error sequences (000, 001,...,111) in three-trial blocks. Analyses of these sequences are reported under the <u>sequence</u> columns; analyses under the <u>count</u> columns combined these sequences to represent differing numbers of errors (with 0 errors in the sequence 000; 1 error in the sequences 001, 010, and 100; 2 errors in the sequences 011, 101, and 110; and 3 errors in the sequence 111).

<u>Line 1</u>. "Gross error" under count is the estimated gross error rate P, the total number of errors divided by the total number of trials.

<u>Line 2</u>. "Moment Est. p,k" under count reports the estimates for p and k in Equation 1 obtained by Samanta & Mitra's moment estimation method (NUREG/CR-2211, 1981). This moment estimation technique analyzes only the frequencies of error counts, despite the model's prediction of different frequencies for sequences containing the same error counts (e.g., 101 and 011), possibly because information on the sequence of errors might be unavailable, or because the sequence of errors is

Table B.2 Within-Block Analysis of All Blocks for Equation 1 Model

	TASK 1				TASK 2				TASK 3				COMPOSITE TASK			
	sequence		count		sequence		count		sequence		count		sequence		count	
	p	k	p	k	p	k	p	k	p	k	p	k	p	k	p	k
S Gross error			.038				.024				.064				.133	
U Moment Est. p,k			.04	.000			.024	.000			.056	.192			.112	.27
B Min. X^2 Est.:																
J Est. p (k=0)	.040		.040		.030		.024		.105		.077		.203		.201	
Residual, X^2=	.127		.127		.983		.051		6.150		1.775		9.130		7.680**	
1 Est. p,k	.040	.000	.040	.000	.030	.000	.024	.000	.065	.465	.057	.283	.132	.335	.129	.284
Residual, X^2=	.127		.127		.983		.051		2.013		.352		3.299		2.218	
Fit for k, X^2=	.000		.000		.000		.000		4.137**		1.423		5.831**		5.462**	
S Gross error			.015				.063				.175				.211	
U Moment Est. p,k			.015	.000			.068	.000			.156	.177			.166	.366
B Min. X^2 Est.:																
J Est. p (k=0)	.026		.015		.088		.068		.191		.188		.253		.245	
Residual, X^2=	1.537		.016		3.648		.299		3.230		2.210		8.067		4.066	
2 Est. p,k	.026	.000	.015	.000	.088	.000	.068	.000	.18	.146	.163	.267	.203	.320	.166	.367
Residual, X^2=	1.537		.016		3.648		.299		2.709		1.354		4.673		.000	
Fit for k, X^2=	.000		.000		.000		.000		.521		.856		3.394*		4.066**	
S Gross error			.042				.049				.094				.184	
U Moment Est. p,k			.043	.000			.051	.000			.108	.000			.171	.118
B Min. X^2 Est.:																
J Est. p (k=0)	.058		.043		.058		.051		.106		.103		.211		.194	
Residual, X^2=	3.382		.185		1.877		.278		1.833		1.086		4.918		1.932	
3 Est. p,k	.058	.000	.043	.000	.058	.000	.051	.000	.106	.000	.103	.000	.206	.064	.176	.183
Residual, X^2=	3.382		.185		1.877		.278		1.833		1.086		4.734		1.347	
Fit for k, X^2=	.000		.000		.000		.000		.000		.000		.184		.585	
S Gross error			.071				.192				.286				.431	
U Moment Est. p,k			.077	.000			.119	.74			.218	.445			.308	.638
B Min. X^2 Est.:																
J Est. p (k=0)	.105		.077		.296		.296		.332		.329		.467		.471	
Residual, X^2=	7.151.516		22.351***		22.351***		11.413*		8.989**		17.805***		13.156***			
4 Est. p,k	.105	.000	.077	.000	.155	.513	.147	.612	.262	.316	.223	.429	.346	.524	.334	.542
Residual, X^2=	7.151		.516		4.268		3.237*		4.904		.512		6.080		3.067*	
Fit for k, X^2=	.000		.000		18.083***		19.114***		6.509**		8.477***		11.725***		10.089***	
S Gross error			.213				.385				.347				.667	
U Moment Est. p,k			no solution				.327	.309			.347	.000			no solution	
B Min. X^2 Est.:																
J Est. p (k=0)	.242		.218		.394		.394		.345		.347		.632		.653	
Residual, X^2=	7.118		.275		7.343		2.804		5.591		.000		12.655**		1.263	
5 Est. p,k	.233	.076	.218	.000	.322	.351	.327	.305	.345	.000	.347	.000	.559	.408	.653	.000
Residual, X^2=	6.930		.275		2.605		.024		5.591		.000		8.906		1.263	
Fit for k, X^2=	.188		.000		4.738**		2.780*		.000		.000		3.749*		.000	

irrelevant for the intended application. (The moment estimation technique was not used in any of the subsequent analyses, and this line is omitted in subsequent tables.)

Remaining entries for a given subject-task combination in the Table describe results of estimating model parameters by minimization of Chi Square (X^2). This technique permits analyses of either sequences or counts, and permits at least nominal tests of significance both for lack of fit and for variation attributable to the parameter k (see Grant, 1962, for a discussion of these two approaches to testing models). While the analyses of sequences provide more searching tests of the model, the analyses of counts are based on larger combined frequencies and should better approximate the theoretical statistical distribution. But for both analyses, the nominal levels of significance should be considered suggestive, since results based on the small expected frequencies may not closely approximate the X^2 distribution. These analyses searched successively finer grids of trial parameter estimates for p and k to find estimates which minimized X^2, using double-precision arithmetic, and values of .00001 and .99999 as approximations of 0 and 1.0 for p and k (in order to avoid zero expected frequencies as denominators).

<u>Line 4</u>. "Est. p (k=0)" reports the minimum X^2 estimate of p assuming independent errors (k=0), for the analysis of error sequences and error counts.

<u>Line 5</u>. "Residual,X^2" is the calculated X^2 associated with the preceding estimate of p, and tests the discrepancy from frequencies expected if errors were independent. This X^2 for lack of fit has 6 d.f. for sequences and 2 d.f. for counts. Nominal levels of significance are indicated by * (.10), ** (.05), and *** (.01).

<u>Line 6</u>. "Est. p,k" gives the minimum X^2 estimates of p and k in the Equation 1 model for sequences and for counts.

<u>Line 7</u>. "Residual,X^2" reports calculated X^2 values associated with the estimates on the preceding line, testing departures from the model or lack of fit. These X^2s have 5 d.f. for the test of sequences and 1 d.f. for the test of counts, with nominal significance indicated as above.

<u>Line 8</u>. "Fit for k,X^2" reports the reduction in X^2 (Line 5 minus Line 7) attributable to fitting k as well as p, and tests the improvement of fit achieved with k. These X^2 tests have 1 d.f. for both sequences and counts, with nominal significance indicated as above.

Departures from independence achieved nominal significance in Table B.2 (Line 5) only for Subject 4 on T2, T3, and CT; and for Subject 1's counts and Subject 5's sequences on CT. All had patterns of p_1, p_2, and p_3 estimates in Figure B.2 which suggested dependence. The tests did not approach significance for the six remaining subject-task combinations where relative estimates of p_1 and p_2 in Figure B.2 suggested departures from independence, nor for the nine cases of apparent independence or possible negative dependence. In

summary, these analyses suggest dependence only when either Subject 4 or the Composite Task was involved.

While there is close agreement between the moment estimation (Line 2) and minimum X^2 estimation (Line 5) of base error rate p based on counts, there is less agreement between their estimates of the dependence parameter k in Samanta & Mitra's model.

3.5 Tests of Indendence Within All Blocks Pooled over Subjects or Tasks

Observed frequencies of error sequences were pooled over subjects or over the three separate tasks, and the X^2 test of independence performed for each pooled data set as in the preceding analyses of data for separate subject-task combinations. While the larger expected frequencies should improve approximation of the theoretical X^2 distribution, some expected frequencies were still small. Furthermore, mixing data with different base error rates could lead to rejection of the hypothesis of independence for the pooled data, even if independence held in each subject-task combination. Even under such unfavorable conditions, the hypothesis of independence could be rejected ($p < .05$) only for Subject 4 and for Subject 5's sequences for data pooled over the three tasks; for the Composite Task with pooled data for Subjects 1-3's counts; for Tasks 2, 3, and the Composite Task when results from Subjects 1-4 were pooled; and for Task 2's sequences and the Composite Task when all five subjects' data were combined. These results are consistent with previous indications of dependence when Subject 4 or the Composite Task were involved, but also suggest possible dependence for Subject 5.

3.6 Fit of the Equation 1 Model Within All Blocks

For these pooled data as well as the individual subject-task analyses in Table B.2, Samanta & Mitra's model successfully accomodated most nominally significant or near-significant departures from independence, with possible rejection of goodness-of-fit of the model (Line 7) in four cases of error counts at the .10 level (for S4-T2, S4-CT, Subject 4's pooled tasks, and Subjects 1-4 pooled for Task 2), and for Subject 5's sequences pooled over tasks at the .05 level.

3.7 Fit of the Equation 1 Model Within Edited Blocks

According to any of the models, when an error occurs the probability of a successive error on the same task is greater than p unless the errors are independent. Dependence might carry over from one block to the next on the same task, with the error probability p_i greater than p_1 for the first trial in a block despite the interpolated task(s), or might be attenuated (as the earlier analyses of sequences over blocks suggested). To eliminate possible carry-over which would distort parameter estimates and goodness-of-fit tests, a block was deleted if an error occurred on the last trial of that task's previous block in the session. The observed frequencies of the eight possible

error sequences for these edited three-trial blocks were tabulated for each task and subject and are reported in Table B.8 in the Attachment.[13]

The edited data included multiple errors in Composite Task blocks for all subjects. Blocks containing multiple errors on the three separate tasks were sparse. For Subjects 1-3 the only four blocks with multiple errors occurred for Subjects 1 and 2 on Task 3. Subject 4's data included a total of four multiple error blocks for Tasks 2 and 3, and Subject 5's data contained 16 multiple error blocks spread over all three tasks.

The results in Table B.3 for edited blocks follow the same format as Table B.2 (except that the results of moment estimation are omitted).

Nominally significant departures from independence are present for the same five subject-task combinations as in Table B.2, although here the analyses of Subject 4's sequences do not approach significance, and the counts for S4-CT are only marginally significant. Where there is such evidence of dependence (now Line 4), except for S5-CT the minimization of X^2 yielded estimates of p=.105-.275 and k=.324-.382 for Samanta & Mitra's model (Line 5). Estimates for S5-CT's sequences were higher with p=.529 and k=.721 (but again with k=0 for the analysis of counts). Except for S5-CT, departures from independence were detected in the analysis of counts rather than sequences, and these four are the only cases where the count of triple errors was higher than the frequency of double errors, in the form of a truncated U-shaped distribution (see attached Table B.8).

For all these cases including S5-CT the improvement in fit attributable to k is nominally significant or near significant (Line 7), and nowhere in Table B.3 do departures from fit of Samanta & Mitra's model approach significance (Line 6) except for S5-CT. Even where dependence was not significant there was at least marginally significant improvement in fit attributable to k for S1-T3's sequences and S2-CT with values of p=.055-.230 and k=.313-.471; and in the sequences for S1-CT and Subject 4 on T2, T3, and CT.

In all these cases the relative estimates of p_1 and p_2 in Figure B.2 had suggested dependence. Estimates of k were in the range 0-.485 for other subject-task combinations where Figure B.2 suggested dependence, and 0-.034 where Figure B.2 did not suggest positive dependence.

---

13. For Subjects 1-3 editing removed 12 blocks on Tasks 1-3 (three with single errors) and 10 on the Composite Task (three with single errors and one double error). For Subject 4 editing removed 13 blocks on Tasks 1-3 (six single, two double, and three triple error blocks), and 8 blocks on the Composite Task (two single, one double, and four triple errors). For Subject 5 editing removed 17 blocks on Tasks 1-3 (six single, two double, and one triple error), and 14 blocks on the Composite task (two single, four double, and four triple errors). Thus, editing reduced the number of blocks with multiple errors, and some of the evidence of dependence for Subjects 4 and 5. The results for Subjects 1 and 2 on Task 1 remain the same as in Table B.2.

Table B.3 Within-Block Analysis of Edited Blocks for Equation 1 Model

	TASK 1				TASK 2				TASK 3				COMPOSITE TASK			
	sequence		count		sequence		count		sequence		count		sequence		count	
	p	k	p	k	p	k	p	k	p	k	p	k	p	k	p	k
S Gross error							.025				.056				.130	
U Min. X^2 Est.:																
B Est. p (k=0)					.031		.025		.104		.076		.206		.205	
J Residual, X^2=					.986		.053		6.842		2.751		8.896		8.225**	
Est. p,k					.031	.000	.025	.000	.055	.471	.047	.421	.125	.340	.123	.324
1 Residual, X^2=					.986		.053		1.829		.526		2.475		1.984	
Fit for k, X^2=					.000		.000		5.013**		2.225		6.421**		6.241**	
S Gross error							.070				.167				.244	
U Min. X^2 Est.:																
B Est. p (k=0)					.097		.075		.191		.191		.282		.275	
J Residual, X^2=					3.737		.337		4.090		4.090		6.421		3.483	
Est. p,k					.097	.000	.075	.000	.177	.148	.148	.485	.230	.313	.190	.403
2 Residual, X^2=					3.737		.337		3.466		2.271		3.606		.013	
Fit for k, X^2=					.000		.000		.624		1.819		2.815*		3.470*	
S Gross error			.043				.051				.100				.173	
U Min. X^2 Est.:																
B Est. p (k=0)	.059		.045		.060		.053		.113		.110		.202		.181	
J Residual, X^2=	3.397		.191		1.891		.287		1.937		1.181		4.139		1.237	
Est. p,k	.059	.000	.045	.000	.060	.000	.053	.000	.113	.000	.110	.000	.202	.000	.167	.155
3 Residual, X^2=	3.397		.191		1.891		.287		1.937		1.181		4.139		.848	
Fit for k, X^2=	.000		.000		.000		.000		.000		.000		.000		.389	
S Gross error			.028				.116				.242				.312	
U Min. X^2 Est.:																
B Est. p (k=0)	.047		.040		.199		.198		.304		.302		.364		.364	
J Residual, X^2=	3.204		.301		10.032		9.804***		8.899		7.603**		10.133		5.110*	
Est. p,k	.047[a]	.002[a]	.040	.000	.108	.352	.105	.382	.227	.324	.208	.355	.281	.392	.275	.325
4 Residual, X^2=	3.204		.301		2.157		1.714		3.517		1.806		5.864[b]		1.916	
Fit for k, X^2=	.000		.000		7.875***		8.090***		5.382**		5.797**		4.269**		3.194*	
S Gross error			.238				.386				.386				.667	
U Min. X^2 Est.:																
B Est. p (k=0)	.269		.244		.400		.396		.382		.386		.633		.659	
J Residual, X^2=	6.958		.320		5.732		1.097		3.541		.411		17.572***		1.879	
Est. p,k	.257	.103	.244	.000	.348	.249	.349	.198	.377	.034	.382	.025	.529	.721	.659	.000
5 Residual, X^2=	6.702		.320		3.927		.166		3.510		.400		10.515*		1.879	
Fit for k, X^2=	.256		.000		1.805		.931		.031		.011		7.057***		.000	

a. Not unique.

b. X^2 residual larger than for the Equation 2 two state model (see text).

There is generally fair agreement between the parameter estimates based on sequences and those based on counts, but with some large discrepancies such as those for S2-T3 and S5-CT, and for S3-CT and S5-T1.

In these analyses Samanta & Mitra's model in Equation 1 is remarkably successful in accomodating the data, both in terms of non-significant residuals and accounting for observed variation. In fact, in Table B.3 where either departure from independence or fit of the model was at least marginally significant, the residual X^2 was too small to permit significance (.05) of any additional parameter in all but two cases (the sequences for Subjects 4 and 5 on CT). Further, the estimated values of k mostly fall in a narrow range where there is any indication of statistical significance. However, as the attached Table B.8 indicates, the analyses were based on small frequencies, and for that reason cannot be considered conclusive. Therefore, several methods of combining analyses over subjects or tasks were attempted.

3.8 Fit of the Equation 1 Model Within Edited Blocks over Subjects or Tasks

Several methods of combining individual subject-task analyses were possible. First, we pooled observed frequencies of error sequences over subjects or over separate tasks, then fitted parameters and calculated X^2s for each pooled data set as in the analyses of data from each different subject-task combination. While employing larger expected frequencies and therefore a better approximation to the theoretical X^2 distribution, the expected frequencies are still small and furthermore the results of these analyses are based on data which presumably represent mixtures of different base error rates and/or degrees of dependence. (However, such pooled data could directly address issues concerning dependence in domains broader than individual subject-task combinations.)

We also pooled the previous separate subject-task tests (summing X^2s and their degrees of freedom from Table B.3) based on separate parameter estimates for each subject-task combination, permitting tests over subjects or over the separate tasks, but still with only nominal significance levels because of the small expected frequencies involved.

Given the complementary shortcomings of these methods, convergence of results from both analyses of the same data would indicate that the conclusions could not be attributed solely to the methodological problem of mixing heterogeneous data nor to the statistical problem of small expected frequencies. Table B.4 summarizes the results of these two approaches to combining results over subjects or tasks, with the columns in Table B.4 corresponding to Lines 4-7 in the analyses in Table B.3. Simultaneous agreement on rejecting the hypothesis of independence on the basis of residuals after fitting p, failing to reject the Equation 1 model on the basis of residuals after fitting k and p, and significant variation attributable to k is a stringent criterion for convergence. On this basis, convergent evidence in Table B.4 supports the conclusion of dependence on the Composite Task (for Subjects 1-3, 1-4, and 1-5); on Task 3 for Subjects 1-4 but not Subjects 1-3 or 1-5; and for Subject 4. Less conservative criteria might find support for dependence on Task 2 for Subjects 1-4 and 1-5, and on Task 3 for Subjects 1-5. These results are

Table B.4 Within-Block Analysis of Edited Blocks for Equation 1 Model Combined over Subjects or Tasks

		POOLED ERROR SEQUENCES				SUMMED X^2			
DOMAIN	DATA	Residual p(k=0)	Est. k	Residual p,k	Fit for k	Residual p(k=0)	Est. k (range)	Residual p,k	Fit for k
TASK 1									
S1-5	seq.	*	.174	ns	*	ns	0-.103	ns	ns
	count	ns	.104	ns	ns	ns	0	ns	ns
S1-4	seq.	ns	0	ns	ns	ns	0-.002	ns	ns
	count	ns	0	ns	ns	ns	0	ns	ns
S1-3	seq.	ns	0	ns	ns	ns	0	ns	ns
	count	ns	0	ns	ns	ns	0	ns	ns
TASK 2									
S1-5	seq.	***	.318	ns	***	ns	0-.352	ns	*
	count	***	.301	ns	***	ns	0-.382	ns	ns
S1-4	seq.	**	.194	ns	***	ns	0-.352	ns	*
	count	***	.189	ns	***	ns	0-.382	ns	*
S1-3	seq.	ns	0	ns	ns	ns	0	ns	ns
	count	ns	0	ns	ns	ns	0	ns	ns
TASK 3									
S1-5	seq.	**	.192	ns	***	ns	0-.471	ns	*
	count	***	.268	ns	***	ns	0-.485	ns	*
S1-4	seq.	*	.192	ns	***	ns	0-.471	ns	**
	count	***	.241	ns	***	**	0-.485	ns	**
S1-3	seq.	ns	.132	ns	ns	ns	0-.471	ns	ns
	count	ns	.179	ns	ns	ns	0-.485	ns	ns
COMPOSITE TASK									
S1-5	seq.	***	.371	ns	***	**	0-.721	ns	***
	count	***	.349	ns	***	**	0-.403	ns	**
S1-4	seq.	***	.266	ns	***	ns	0-.392	ns	***
	count	***	.287	ns	***	**	.155-.403	ns	***
S1-3	seq.	*	.186	ns	**	ns	0-.340	ns	**
	count	**	.280	ns	***	**	.155-.403	ns	**
TASKS 1-3									
S1	seq.	ns	.232	ns	**	ns	0-.471	ns	ns
	count	ns	.132	ns	ns	ns	0-.421	ns	ns
S2	seq.	ns	.154	ns	ns	ns	0-.148	ns	ns
	count	ns	.294	ns	*	ns	0-.485	ns	ns
S3	seq.	ns	0	ns	ns	ns	0	ns	ns
	count	ns	0	ns	ns	ns	0	ns	ns
S4	seq.	***	.380	ns	***	ns	.002-.352	ns	***
	count	***	.382	ns	***	***	0-.382	ns	**
S5	seq.	*	.161	*	ns	ns	.034-.249	ns	ns
	count	ns	.093	ns	ns	ns	0-.198	ns	ns

Key for significance levels: not significant (ns), .10 (*), .05 (**), .01 (***)

consistent with the earlier indications of dependence when the Composite Task, Subject 4, or possibly Subject 5 was involved. Table B.4 also shows how well Samanta & Mitra's model accomodated any departures from independence, with only a single significant residual after fitting p and k.

Finally, differences in parameter estimates between subjects and between tasks were tested by analyses of variance of individual subject-task parameter estimates in the 3 Tasks x 3-5 Subjects design. This approach avoids estimates based on pooling possibly heterogeneous data, avoids the issue of the contribution of small expected frequencies to individual subject-task X^2 values, and is often considered reasonably robust in the face of varying underlying distributions. This approach also tests differences directly, rather than inferring differences indirectly (and possibly improperly) from different levels of significance. While it is not possible to test the overall hypothesis k=0 because estimates of k cannot be negative, significant differences in k imply dependence in some cases.[14] Unfortunately, tests for such a two-factor design with one observation per cell are relatively insensitive, and increasingly conservative if subject-task interactions are present. Furthermore, it is not obvious how the parameters should be scaled, offering the option of testing either the parameter estimates or logarithms of their estimates (plus .001). One shortcoming of these analyses is that they do not include the Composite Task.

In the analyses of base error rate or its logarithm, for parameter estimates based on sequences or for counts, all four analyses found significant differences between subjects and between tasks in analyses of Subjects 1-5, significant differences between tasks for Subjects 1-4, and no significant differences in analyses for Subjects 1-3, suggesting that such analyses might be sensitive enough to detect differences in dependence also. In the analyses of Subjects 1-4 the proportion of observed variance accounted for by tasks ($\underline{eta}^2 = SS_{TASKS}/SS_{TOTAL}$) was .49-.50, with .23-.28 of the variance attributable to subjects.

In the 12 corresponding analyses of estimated values of Equation 1's dependence parameter k or its logarithm, for sequences or for counts, and for 3, 4, or 5 subjects, the only marginally significant (.10) effects were for tasks in the analysis of log k based on counts for Subjects 1-5, and in three of four analyses for Subjects 1-4 (excepting the analysis of k based on sequences). This is some indication of different levels of dependence between tasks, and therefore of dependence on some tasks. The finding lends some support to results of the earlier combined analyses suggesting differences between tasks, but not the suggested differences between subjects. In these

14. Potentially the most powerful test of dependence, that the mean of k is not zero over all conditions, cannot be used because estimates of k cannot be negative in any of these dependence models. Considering only positive dependence (in accordance with a conservative policy of not underestimating risk) may have the ironic effect of making it more difficult to detect positive dependence. For purposes of research, a restructuring of models and reparameterization to permit negative dependence might be useful.

four analyses the eta^2 for tasks was .39-.45, with .20-.33 of the variance attributable to subjects. Compared with the analysis of error rate, there may be relatively less stable parameter estimates, smaller differences between subjects or tasks, or larger interactions between subjects and tasks for the parameter k.

Given this inconclusive result, we returned to the direct estimates of p_1 and p_2 represented in Figures B.1 and B.2 to test for differences in dependence between subjects and/or the separate tasks, using the same analysis of variance design as above. For the Equation 1 and Equation 4 models $k = (p_2 - p_1)/(1 - p_1)$; for the Equation 2 model $k = (p_2 - p_1)$; and for the Equation 3 model $k = (p_2/p_1)$. Such estimates of k calculated from estimated p_1 and p_2 values could take on values representing negative dependence; therefore corresponding variables were defined with estimates representing negative dependence replaced by the value of k representing independence. In addition logarithms (plus .001) of each of the preceding variables which had the form a ratio were analyzed (for those ratios which could not take on negative values). Thus there were 54 analyses of variance of the subjects x tasks factorial layout, for the 9 variables (above) x 2 data sets (Figure B.1 or Figure B.2 data) x 3 subsets of subjects (1-3, 1-4, or 1-5).

Some of these analyses permitted a test of independence based on the mean of all cases. Although these analyses do not involve the crucial estimate of p_3, if subject-task dependence can be described by an additive model for subject and task parameters, then for the appropriate measure the between-subject and between-task differences should be maximized, while their interactions and the error term should be minimized since they would not include errors of scale, thus yielding greater significance and variance accounted for.

In all six analyses for different data sets and subjects, the overall mean of log(k+.001) for the Equation 3 model was significantly less than zero, representing negative dependence. This anomaly was apparently produced by the log transformation of the Equation 3 measure when there were estimates of $p_2 = 0$, and will be disregarded. For both this measure and for log(k+.001) of the Equation 1 measure (with non-negative estimates) of k, and for both data sets, the effect of tasks was marginally significant in the analysis of Subjects 1-4. For both measures the eta^2 for tasks was .45-.49 for Figure B.1 data while eta^2 for subjects was .22-.23, with corresponding eta^{2}s of .33 and .19-.22 for Figure B.2 data.[15] The only other marginally significant effects in all these analyses were differences between tasks for Figure B.1 data, again for Equation 3's log(k+.001) for Subjects 1-5, and for Equation 2's k (with non-negative estimates) for Subjects 1-4. In summary, for Subjects 1-4 roughly the same proportions of variance are attributable to tasks and to subjects as in the previous analysis of minimum X^2 estimates, with similar

---

15. For purposes of comparison with minimum X^2 estimates, the obtained values of k on its original scale for Tasks 1, 2, and 3 are 0, .002, and .049-.057, respectively, for the Model 1 measure, with corresponding values of 0, .005, and .355-.412 for the Model 3 measure.

marginally significant effects. Also, whether k is defined by the Equation 1 or Equation 3 model, the logarithms of either estimate yield about the same results.

From these combined analyses we conclude that dependence is present in some cases, on the basis of the convergent evidence, with more confidence than from the results for separate subject-task combinations. Such apparent dependence is consistent with results for the individual subjects and tasks, and Samanta & Mitra's model successfully accomodated such departures from independence. However, evidence for differences in dependence between subjects and/or tasks is either indirect or only marginally significant.

3.9 Fit of Alternative Models for Edited Blocks

The following analyses not only compare several models, but, since any mechanism which permitted increasing p_2 and p_3 might improve a model's fit for observed multiple errors, may also suggest whether a given improvement in fit is impressive or routine. We already noted in Table B.3 that the two-state model fit S4-CT's sequences better than the Equation 1 model. Table B.5 compares the residual X^2 after fitting alternative models, and gives details of fit for the positively accelerated Equation 3 model (with estimates of p and k on Line 5 as defined in Equation 3), for those cases where either departure from independence or the Equation 1 model's fit was near significant in Table B.3.[16] One consideration in the Equation 2 and 3 models is that the "min" term provides additional latitude in fitting, which should be recognized and discounted if without the "min" term the model would predict probabilities greater than one. There were only a few cases, indicated by a superscript "a" for the estimated k for Model 3, where the fit might be suspicious by reason of this latitude. Discounting such fit, Samanta & Mitra's Equation 1 model best fits two cases (S1-T3's sequences and S2-CT's counts), Equation 2 fits one (S5-CT's sequences), the two-state model fits one (S4-CT's sequences), and Equation 3 best fits the remaining eight. For six of these last eight cases, the best fitting k yielded values of p_3 close to one. This may suggest an even more positively accelerated increase than provided by Equation 3 (since a less positively accelerated increase would have been fitted with $p(1+k)^2$ greater than one, i.e., with p_3 greater than one except for the "min" term).

These comparisons indicate only relative fit, but certainly not statistically significant differences in fit, and might depend on the particular estimation technique. In many cases the differences were quite small, and these results (based on somewhat different data) do not agree with the apparent fit of models in Figures B.1 and B.2. At the very least these comparisons of fit make Equation 1's success in accommodating data less impressive.

---

16. Note that the analyses of counts for S1-T3 and S5-CT are not included in this discussion, because of failure to achieve near-significance in the earlier analysis, although they are reported in Table B.5.

Table B.5 Within-Block Analysis of Edited Blocks for Equation 3 Model

		TASK 1				TASK 2				TASK 3				COMPOSITE TASK			
		sequence		count		sequence		count		sequence		count		sequence		count	
		p	k	p	k	p	k	p	k	p	k	p	k	p	k	p	k
S	Gross error											.056				.130	
U	Min. X^2 Est.:																
B	Est. p (k=0)									.104		.076		.206		.205	
J	Residual, X^2=									6.842		2.751		8.896		8.225**	
	Est. p,k									.055	8.09[a]	.049	11.02[a]	.119	1.90[b]	.115	1.95[b]
1	Residual, X^2=									1.829		.847[c]		1.739		1.142	
	Fit for k, X^2=									5.013**		1.904		7.157***		7.083***	
S	Gross error															.244	
U	Min. X^2 Est.:																
B	Est. p (k=0)													.282		.275	
J	Residual, X^2=													6.421		3.483	
	Est. p,k													.224	1.08	.203	1.00
2	Residual, X^2=													3.267		.269[c]	
	Fit for k, X^2=													3.154*		3.214*	
S	Gross error																
U	Min. X^2 Est.:																
B	Est. p (k=0)																
J	Residual, X^2=																
	Est. p,k																
3	Residual, X^2=																
	Fit for k, X^2=																
S	Gross error							.116				.242				.312	
U	Min. X^2 Est.:																
B	Est. p (k=0)					.199		.198		.304		.302		.364		.364	
J	Residual, X^2=					10.032		9.304***		8.899		7.603**		10.133		5.110*	
	Est. p,k					.102	2.13[b]	.101	2.15[b]	.214	1.16[b]	.196	1.26[b]	.296	.56	.256	.96
4	Residual, X^2=					1.289		.937		2.534		.800		6.388[c]		1.160	
	Fit for k, X^2=					8.743***		8.867***		6.365**		6.803***		3.745*		3.950**	
S	Gross error															.667	
U	Min. X^2 Est.:																
B	Est. p (k=0)													.633		.659	
J	Residual, X^2=													17.572***		1.879	
	Est. p,k													.529	.64[a]	.659	.00
5	Residual, X^2=													10.484*[d]		1.879	
	Fit for k, X^2=													7.088***		.000	

a. Estimated $p(1+k)^2$ greater than 1.
b. Estimated $p(1+k)^2$ approximately 1.
c. X^2 residual larger than for the Equation 1 model.
d. X^2 residual larger than for the Equation 2 linear model.

3.10 Fit of Alternative Models for Edited Blocks over Subjects or Tasks

Where the earlier analyses of pooled responses indicated (near) significant departure from independence or improvement in fit attributable to k in the Equation 1 model, the residual X^2s after fitting the Equation 1 and Equation 3 models were compared. For each subject's sequences and counts pooled over the three tasks, Samanta & Mitra's model fit better for Subjects 1, 2, and 5, but worse for Subject 4. For each of the four tasks pooled over Subjects 1-5, Samanta & Mitra's model fit better. However, for the pooled responses of Subjects 1-4 (on Task 2, Task 3, and the Composite Task), and for Subjects 1-3 on the Composite Task, the Equation 3 model had smaller residuals.

3.11 Presumed Effects of Trial Position Within Block on Relative Fit and Parameter Estimates

The previously suggested effect of trial position within block, based on estimates of p_1 where dependence was presumably excluded, could inflate the frequencies of multiple errors and within-block estimates of p_3 above the values they would have without an end or third trial effect. Such effects would increase the estimate of dependence for any of the four models, and affect relative goodness-of-fit of the various models. Formal consideration of such position effects in the minimum X^2 estimates was not possible for the analysis of counts (because of too few degrees of freedom), and was not attempted for the analysis of sequences because of the smaller expected frequencies in those analyses.

Elimination of such a presumed end or third trial effect would improve fit for the earlier models, and decrease fit for the later models in the following sequence: the two-state model (Equation 4), a negatively accelerated increase model (e.g., Samanta & Mitra's Equation 1 model), the linear model (Equation 2), and a positively accelerated increase model (e.g., Equation 3). Thus, the suggested trial position effects make it difficult to draw conclusions from the preceding comparisons of relative fit. The estimated effects of various positions reported in the attached Table B.6 could be used informally for this purpose, recalling that these estimates are based on data different from those used in parameter estimation, and that it is not clear how combined dependency and position effects would affect error rate. On the other hand, adjustments for effects of trial position on dependence should not be made if such effects are generic and triple repetitions are of interest.

4.0 CONCLUSIONS AND DISCUSSION

The separate subject-task analyses and combined analyses provide consistent, convergent evidence of dependence, but largely limited to cases which include either Subject 4 or the artificial Composite Task, and possibly Subject 5, whose error rates were high. Such results justify the study of error dependence at the molar level of complex tasks, avoiding the reduction of molar dependence to molecular analyses of subordinate tasks. However, for many subject-task combinations dependence was not apparent, possibly because of the limited number of observations in the presence of low error rates and

absence of high dependence, or possibly because some circumstance in the experiment reduced dependence. Especially in the analyses of separate subject-task combinations the sensitivity of tests and the correspondence between actual and nominal levels of significance can be questioned because of the small expected frequencies involved.

Evidence for differences in dependence between subjects and/or tasks was either indirect (and possibly incorrect, since differences in significance do not imply significant differences) or only marginally significant. However, tests available for direct comparisons of dependence were not particularly sensitive.

The analyses of responses pooled over subjects or tasks, presumably with closer correspondence between actual and nominal levels of signficance, would also be applicable to domains broader than separate subject-task combinations. There is some rationale for pooling results on a given task over individuals even though their performance varies. Viewing an entire installation as a domain, it seems irrelevant whether error dependence is uniform over all workers, or some individuals contribute more dependent errors, unless it is possible to identify and select workers with less dependence.

Where dependence was apparent, Samanta & Mitra's Equation 1 model successfully accomodated the data both for the separate subject-task combinations and for data pooled over subjects or tasks. However, comparing the residuals after fitting four alternative models, other models often fit the observations slightly better. Unfortunately, the conclusions from these comparisons are ambiguous because of possible confounded effects of different trial positions. The results may also depend on the particular estimation technique employed. At the very least the fit of the Equation 1 model does not seem impressive when compared with other models, and the results here do not identify a leading candidate for the form of dependence under the conditions of this experiment.

As a practical matter, theoretical studies of the sensitivity of risk to alternate possible forms of dependence among human errors should be carried out in order to determine whether this is a question worth answering. Assuming that form of dependence did affect risk, the possible alternative dependence functions should discourage attempts to use minimal data to assess dependence, with the burden of information extraction carried by the dependence model's assumptions. If possible the data, regardless of its source, should permit evaluation of any assumed form of dependency. Expert judgements, for example, should be able to provide the information necessary to test assumptions of the model or specify the appropriate model for dependence, as well as estimating its parameters.

Experience with fitting models to slightly varied data suggests that parameter estimates for any model and fitting technique are extremely sensitive to variations in multiple occurences of errors. Therefore, applications of a model should not depend on unreliable data for the rare but crucial multiple error occurences. Direct human judgements of rare events'

frequencies are likely to be unreliable, or might be influenced by calculations based on formal or personal assumptions about probabilities.

While there are numerous differences between the conditions of the experiment and of the applications to which generalization might be desirable, two may be worth noting. Subjects' knowledge that their performance was recorded and evaluated was unavoidable in the experiment, but serves to raise the practical issue of the cost-effectiveness of outside inspectors' spot-checks (by physically checking or repeating test, calibration, and maintenance tasks), structured to provide incentives for improving technicians' performance. The second concerns the temporal spacing of tasks and leads us to entertain a superficially paradoxical hypothesis, given the presumed decrease in dependence with increasing separation of trials. Compared with the scheduled cycle of task repetitions in a nuclear power plant, even assuming that they were repeatedly performed by the same technician, subjects in the experiment had much more concentrated practice on each task. Distinguishing between the within-block and between-block inter-trial intervals in the experiment (or in a nuclear power plant), we conjecture that dependence within blocks (and thus multiple errors) will increase with increasing between-block intervals. This conjecture questions the applicability of results from any experimental design like ours.

The apparent disproportionate frequency of breaks in error sequences between blocks suggested decreased dependence over time or interpolated activities as would generally be expected (although trial position effects may also have contributed to this result). Unfortunately, it was not feasible to include an experimental condition where there were no successive presentations of the same task within cycles including three trials on each of the three tasks. If interpolated tasks sufficiently reduced dependence of errors between repetitions of a task, then such a staggered schedule of tasks would be worth considering in the interests of effective redundancy, even at the cost of decreased productivity and increased base error rates.

ATTACHMENT TO APPENDIX B

Task Descriptions

Task procedures and photographs of the displays are given for Task 1, "Valve Isolation and Check," Task 2, "Wiring and Metering," and Task 3, "Logic Sequence Checks," on the following pages. Below is a brief scenario for each task.

In "Valve Isolation and Check" the valves on the initial display (Figure B.3) must be closed in specified order and then a latch pressed to open the cover, revealing a second display (Figure B.4). Tests are performed on the row of switches in a specified order, with subsequent responses for each position determined by the indicator light above the switch. The duration of presses is restricted, but each response's duration is displayed on the timer on the upper left. If the meter on the lower right drifts outside a specified range it must be adjusted before the next action. Following completion of these steps, the cover is closed to return to the first display (Figure B.3). After opening the valves the restore switch must be pressed, and the task ended.

In "Wiring and Metering" two lights in Figure B.5 in the indicator bank on the upper left (in which all the lights intermittently blink on and off) must be off whenever any action is performed, and can be turned off by pressing switches. After setting the meter function, a specified series of connections must be made between terminals in the wiring area on the right. After connecting each pair of terminals, the connection is tested and reset if the meter registers outside a specified range, before completing the connection. After all the connections are completed the wiring area must be cleared and the task ended.

In "Logic Sequence Tests" a reset switch must first be pressed to change the indicator lights across the display (Figure B.6). Then a specified series of logic blocks must tested by pressing the switch under the lit or unlit indicator while the indicator lights across the entire display change. If the light in the tested block changes then reset must be pressed before going on to the next block. After completion of the tests the task must be ended.

Each of these task displays offered the same "menu" of administrative actions (at the top of the screen). Between two blocks of different tasks the display appeared as in Figure B.7. Between two trials on the same task only the "review" and "resume (next) task" options were available.

TASK PROCEDURE: VALVE ISOLATION AND CHECK

1) Close valve B.

2) Close valve C.

3) Close valve A.

4) Press latch L to open the COVER.

5) Check METER to be sure it is in the range 13.35-14.65.

If METER is above 14.65 press H until it is in range.

If METER is below 13.35 press L until it is in range.

AT ANY TIME DURING THE FOLLOWING CHECKS (through Step 25) THE METER IS OUTSIDE RANGE, BEFORE NEXT ACTION FOLLOW THE ABOVE PROCEDURE TO BRING IT INTO RANGE.

6) Press switch 147 momentarily (less than 1 sec.). Indicator above switch will go on while you are pressing.

7) Press test switch T for at least 1 sec. If indicator above 147 blinks then return to step 6; otherwise proceed to next test.

8) Press switch 033 momentarily (less than 1 sec.). Indicator above switch will go on while you are pressing.

9) Press test switch T for at least 1 sec. If indicator above 033 blinks then return to step 8; otherwise proceed to next test.

10) Press switch 702 momentarily (less than 1 sec.). Indicator above switch will go on while you are pressing.

11) Press test switch T for at least 1 sec. If indicator above 702 blinks then return to step 10; otherwise proceed to next test.

12) Press switch 970 momentarily (less than 1 sec.). Indicator above switch will go on while you are pressing.

13) Press test switch T for at least 1 sec. If indicator above 970 blinks then return to step 12; otherwise proceed to next test.

14) Press switch 884 momentarily (less than 1 sec.). Indicator above switch will go on while you are pressing.

15) Press test switch T for at least 1 sec. If indicator above 884 blinks then return to step 14; otherwise proceed to next test.

16) Press switch 220 for at least 1 sec. but no more than 2 sec. Indicator above switch will go on while you are pressing.

17) Press test switch T for at least 1 sec. If indicator above 220 blinks then return to step 16; otherwise proceed to next test.

18) Press switch 864 momentarily (less than 1 sec.). Indicator above switch will go on while you are pressing.

19) Press test switch T for at least 1 sec. If indicator above 864 blinks then return to step 18; otherwise proceed to next test.

20) Press switch 377 momentarily (less than 1 sec.). Indicator above switch will go on while you are pressing.

21) Press test switch T for at least 1 sec. If indicator above 377 blinks then return to step 20; otherwise proceed to next test.

22) Press switch 201 momentarily (less than 1 sec.). Indicator above switch will go on while you are pressing.

23) Press test switch T for at least 1 sec. If indicator above 201 blinks then return to step 22; otherwise proceed to next test.

24) Press switch 222 for at least 1 sec. but no more than 2 sec. Indicator above switch will go on while you are pressing.

25) Press test switch T for at least 1 sec. If indicator above 222 blinks then return to step 24; otherwise proceed to next step.

26) Press CLOSE to close cover.

27) Press latch L to secure COVER.

28) Open valve A.

29) Open valve C.

30) Open valve B.

31) Press restore switch R.

32) Press END to signify completion of task.

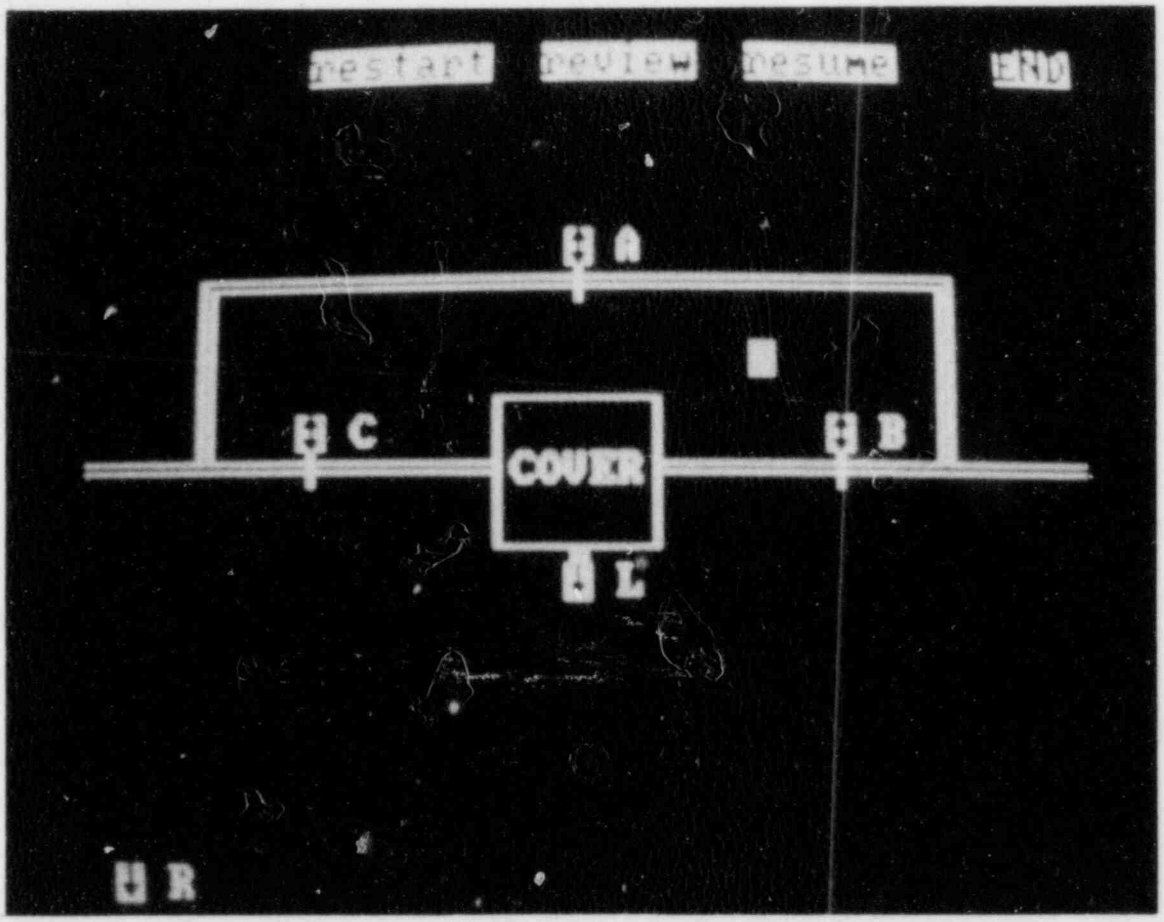


Figure B.3

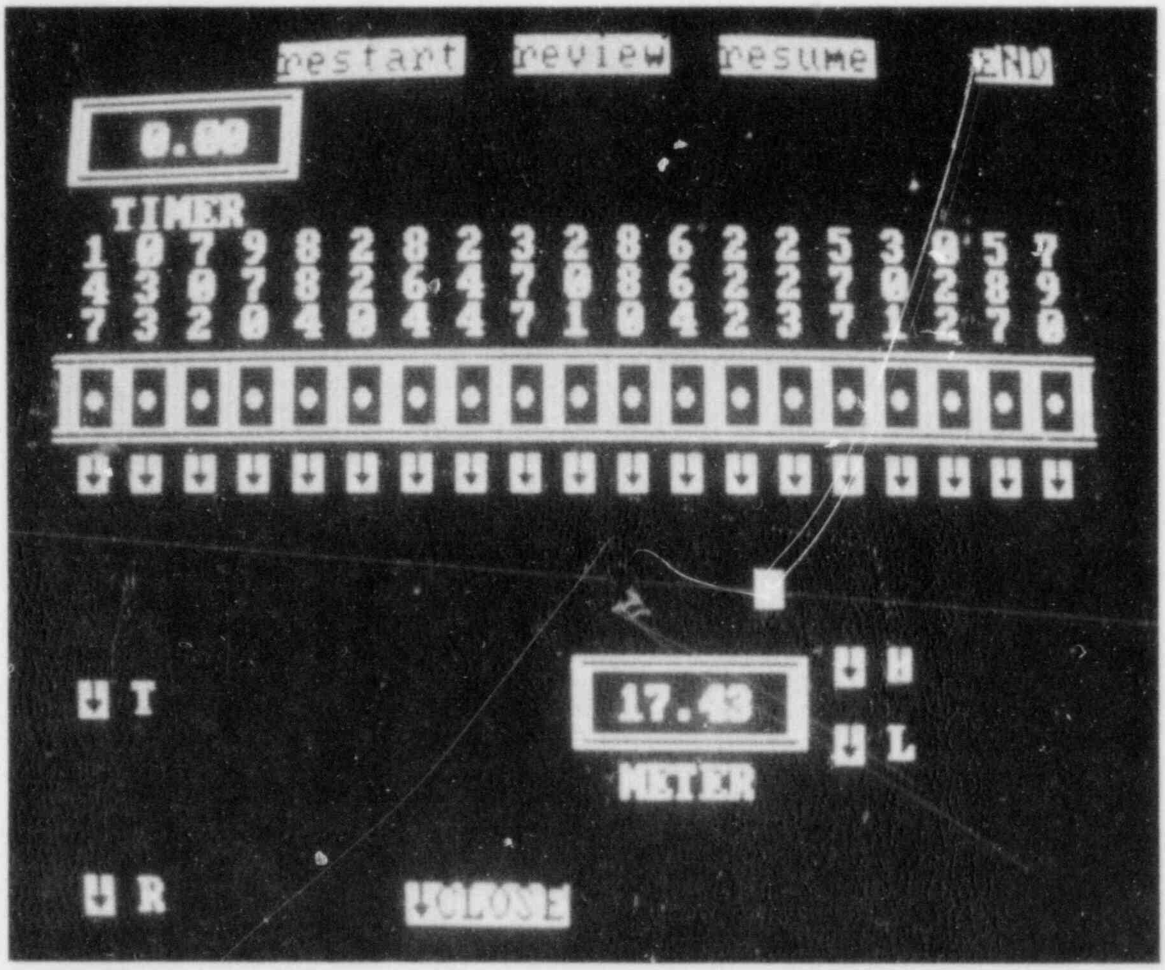


Figure B.4

TASK PROCEDURE: WIRING AND METERING

1) Check to make sure indicators in column K/row H and column Q/row B (in the bank of indicators on the upper left) are off. If K/H and/or Q/B are on, press S2 to turn off K/H; press S1 to turn off Q/B.

AT ANY TIME DURING THE FOLLOWING STEPS (through Step 45) WHEN K/H OR Q/B COMES ON, BEFORE NEXT ACTION FOLLOW THE ABOVE PROCEDURE TO TURN IT OFF.

2) Press O to select that FUNCTION for meter.

3) DO NOT change RANGE setting during this task.

4) DO NOT change METER setting by pressing U or D during this task.
5) Go on to step 6.

6) Connect wire from terminal (next to) 76 ... (by pressing the button with the cursor at the 76 arrow and moving the cursor...)

7) ...to terminal (next to)65.
(...until you press the button with the cursor at the arrow)

8) Press test switch T. If METER registers outside the range 0.50 to 1.00, press reset switch R as many times as necessary to bring METER into this range.
NOTE: the values .50 and 1.00 are within the required range.

9) Press switch C to activate the connection.

10) Connect wire from terminal 67 ...

11) ...to terminal 73.

12) Press test switch T. If METER registers outside the range 0.50 to 1.00, press reset switch R as many times as necessary to bring METER into this range.

13) Press switch C to activate the connection.

14) Connect wire from terminal 84 ...

15) ...to terminal 97.

16) Press test switch T. If METER registers outside the range 0.50 to 1.00, press reset switch R as many times as necessary to bring METER into this range.

17) Press switch C to activate the connection.

18) Connect wire from terminal 95 ...

19) ...to terminal 96.

20) Press test switch T. If METER registers outside the range 0.50 to 1.00, press reset switch R as many times as necessary to bring METER into this range.

21) Press switch C to activate the connection.

22) Connect wire from terminal 94 ...

23) ...to terminal 87.

24) Press test switch T. If METER registers outside the range 0.10 to 0.65, press reset switch R as many times as necessary to bring METER into this range.

25) Press switch C to activate the connection.

26)

27) Connect wire from terminal 86 ...

28) ...to terminal 64.

29) Press test switch T. If METER registers outside the range 0.10 to 0.65, press reset switch R as many times as necessary to bring METER into this range.

30) Press switch C to activate the connection.

31) Connect wire from terminal 93 ...

32) ...to terminal 68.

33) Press test switch T. If METER registers outside the range 0.10 to 0.65, press reset switch R as many times as necessary to bring METER into this range.

34) Press switch C to activate the connection.

35) Connect wire from terminal 69 ...

36) ...to terminal 85.

37) Press test switch T. If METER registers outside the range 0.60 to 0.90, press reset switch R as many times as necessary to bring METER into this range.

38) Press switch C to activate the connection.

39) Go to next step.

40) Go to next step.

41) Go to next step.

42) Connect wire from terminal 63 ...

43) ...to terminal 83.

44) Press test switch T. If METER registers outside the range 0.60 to 0.90, press reset switch R as many times as necessary to bring METER into this range.

45) Press switch C to activate the connection.

46) Clear wiring area (by pressing button while cursor is positioned in the wiring area).

NOTE: IF WIRING AREA IS CLEARED BEFORE COMPLETION OF ALL STEPS THROUGH 46 THEN IT IS NECESSARY TO RESTART THE TASK.

48) Press END to signify completion of task.

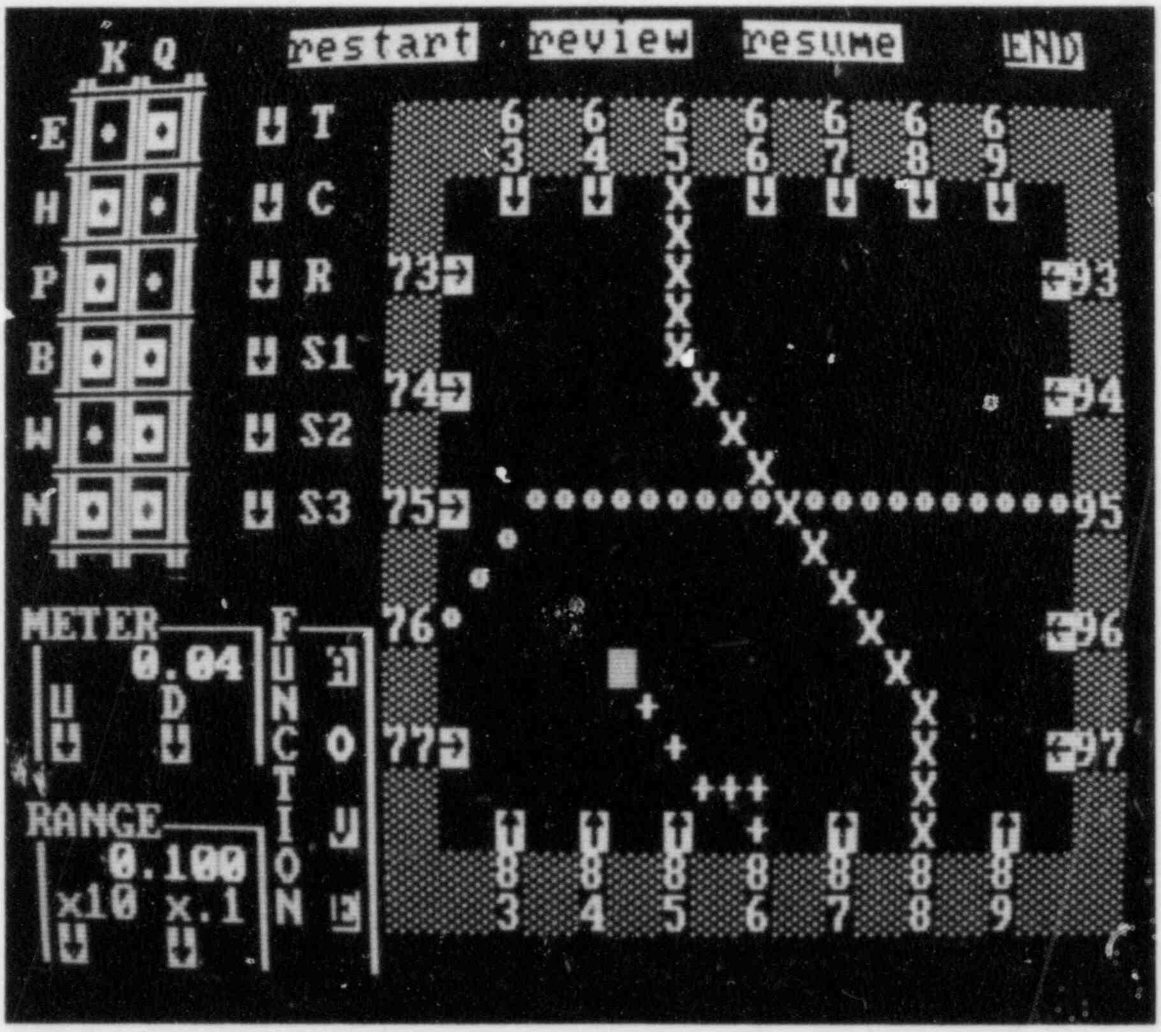

restart
review
resume
END
K Q
E
H
P
B
W
N
T
C
R
S1
S2
S3
METER
0.04
U D
RANGE
0.100
x10 x.1
FUNCTION
63 64 65 66 67 68 69
73 74 75 76 77
93 94 95 96 97
83 84 85 86 87 88 89

Figure B.5

TASK PROCEDURE: LOGIC SEQUENCE TESTS

1) Press reset switch R until scan of lights is completed.

2) In Block C23 press switch under lit position until scan is completed.
If light in block changed, before proceeding press R until scan is completed.

3) In Block C27 press switch under lit position until scan is completed.
If light in block changed, before proceeding press R until scan is completed.

4) In Block C29 press switch under lit position until scan is completed.
If light in block changed, before proceeding press R until scan is completed.

5) In Block C31 press switch under lit position until scan is completed.
If light in block changed, before proceeding press R until scan is completed.

6) In Block G46 press switch under unlit position until scan is completed.
If light in block changed, before proceeding press R until scan is completed.

7) In Block F56 press switch under unlit position until scan is completed.
If light in block changed, before proceeding press R until scan is completed.

8) In Block E56 press switch under unlit position until scan is completed.
If light in block changed, before proceeding press R until scan is completed.

9) In Block G39 press switch under lit position until scan is completed.
If light in block changed, before proceeding press R until scan is completed.

10) In Block G40 press switch under lit position until scan is completed.
If light in block changed, before proceeding press R until scan is completed.

11) In Block G41 press switch under unlit position until scan is completed.
If light in block changed, before proceeding press R until scan is completed.

12) In Block G42 press switch under unlit position until scan is completed.
If light in block changed, before proceeding press R until scan is completed.

13) In Block G43 press switch under lit position until scan is completed.
If light in block changed, before proceeding press R until scan is completed.

14) In Block G44 press switch under lit position until scan is completed.
If light in block changed, before proceeding press R until scan is completed.

15) In Block G45 press switch under unlit position until scan is completed.
If light in block changed, before proceeding press R until scan is completed.

16) In Block F53 press switch under lit position until scan is completed.
If light in block changed, before proceeding press R until scan is completed.

17) In Block C28 press switch under lit position until scan is completed.
If light in block changed, before proceeding press R until scan is completed.

18) In Block E50 press switch under lit position until scan is completed.
If light in block changed, before proceeding press R until scan is completed.

19) In Block C25 press switch under lit position until scan is completed.
If light in block changed, before proceeding press R until scan is completed.

20) In Block G46 press switch under lit position until scan is completed.
If light in block changed, before proceeding press R until scan is completed.

21) In Block F48 press switch under lit position until scan is completed.
If light in block changed, before proceeding press R until scan is completed.

22) In Block F49 press switch under unlit position until scan is completed.
If light in block changed, before proceeding press R until scan is completed.

23) In Block F51 press switch under lit position until scan is completed.
If light in block changed, before proceeding press R until scan is completed.

24) In Block F52 press switch under unlit position until scan is completed.
If light in block changed, before proceeding press R until scan is completed.

25) In Block F54 press switch under lit position until scan is completed.
If light in block changed, before proceeding press R until scan is completed.

26) In Block E55 press switch under unlit position until scan is completed.
If light in block changed, before proceeding press R until scan is completed.

27) In Block D31 press switch under lit position until scan is completed.
If light in block changed, before proceeding press R until scan is completed.

28) In Block D29 press switch under lit position until scan is completed.
If light in block changed, before proceeding press R until scan is completed.

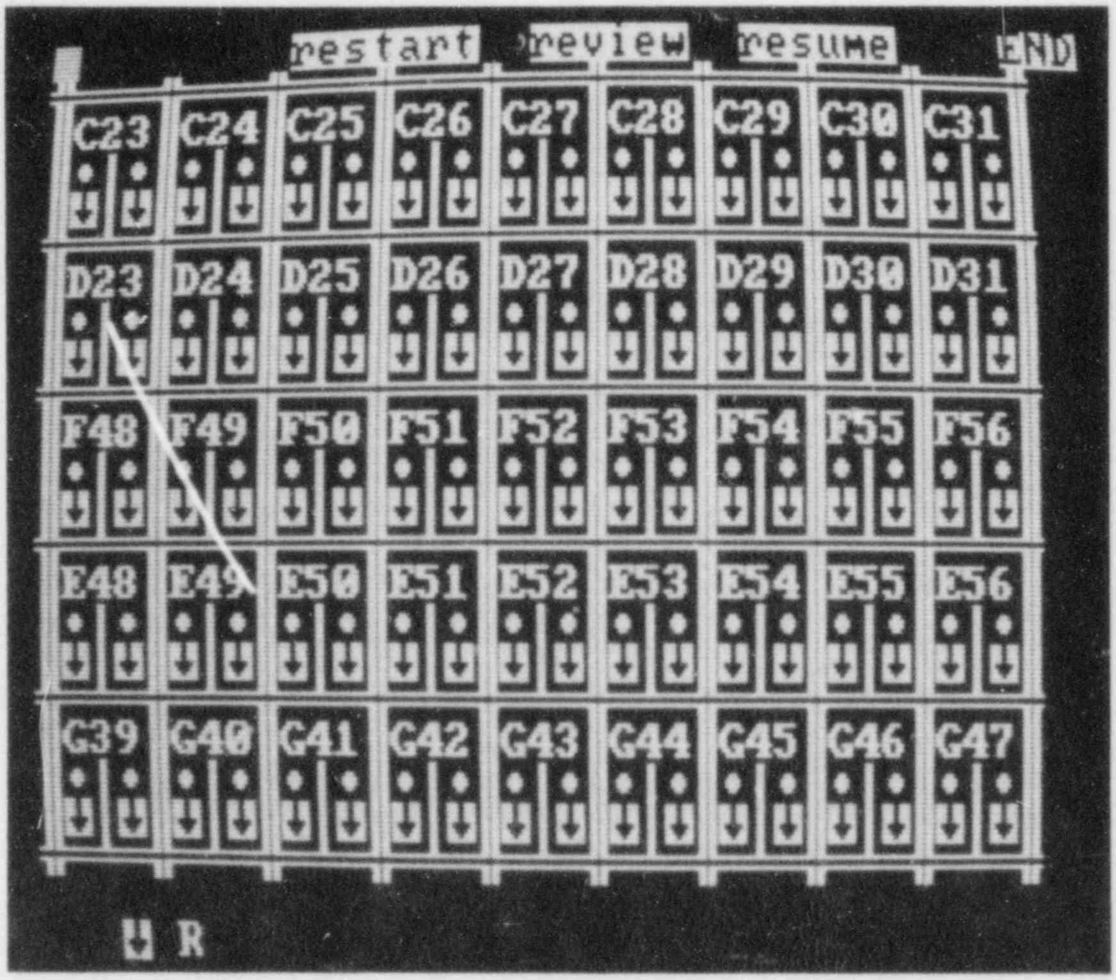
29) In Block D27 press switch under lit position until scan is completed.
If light in block changed, before proceeding press R until scan is completed.

30) In Block D26 press switch under lit position until scan is completed.
If light in block changed, before proceeding press R until scan is completed.

31) In Block D24 press switch under lit position until scan is completed.
If light in block changed, before proceeding press R until scan is completed.

32) In Block D23 press switch under unlit position until scan is completed.
If light in block changed, before proceeding press R until scan is completed.

33) Press END to signify completion of task.



restart
review
resume
END
C23
C24
C25
C26
C27
C28
C29
C30
C31
D23
D24
D25
D26
D27
D28
D29
D30
D31
F48
F49
F50
F51
F52
F53
F54
F55
F56
E48
E49
E50
E51
E52
E53
E54
E55
E56
G39
G40
G41
G42
G43
G44
G45
G46
G47
R

Figure B.6

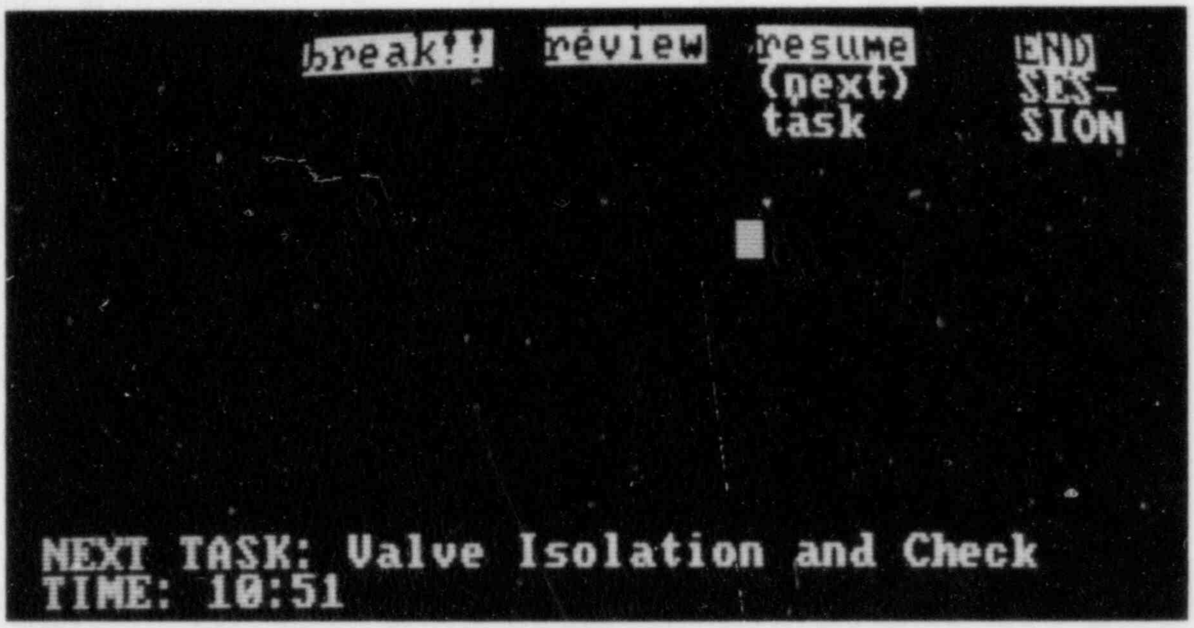

break!!
review
resume
(next)
task
END
SES-
SION
NEXT TASK: Valve Isolation and Check
TIME: 10:51

Figure B.7

Table B.6 Estimated Base Error Rate p for Each Trial Position Within Block

		Trial Position Within Block 1	2	3
Subject 1	Task 1	.038	.040	.040
	Task 2	0.0	.036	.038
	Task 3	.040	.040	.080
	Composite	.087	.087	.182
Subject 2	Task 1	0.0	.045	0.0
	Task 2	.053	0.0	.143
	Task 3	.167	.111	.211
	Composite	.133	.118	.312
Subject 3	Task 1	0.0	.031	.097
	Task 2	.091	.032	.030
	Task 3	.100	.138	.071
	Composite	.200	.208	.167
Subject 4	Task 1	0.0	.036	.185
	Task 2	.130	.095	.095
	Task 3	.182	.095	.318
	Composite	.250	.125	.438
Subject 5	Task 1	.048	.280	.294
	Task 2	.250	.263	.438
	Task 3	.211	.400	.438
	Composite	.273	.714	.778

Table B.7 Sequences of Errors in Task Performance

S1-T1	000 000/000 000 000/000 000 000/000 100 000/000 000 000/000 000 000/000 000 000/000 010 000/000 000 001/
-T2	000 000/001 000 000/000 000 000/000 000 000/000 000 000/000 000 000 000/000 000 000 000/000 000 000/010 000 000/
-T3	000 001 000/000 000 000/000 000 100 0--/000 011 001/000 000 000/000 000 000/ 000 000 000/000 000/000 000 000 ?--/
-CT	000 001/001 000 000/000 000 100/000 111 001/000 000 000/000 000 000/000 000 000/000 010/010 000 001/
S2-T1	000 000 000/000 000 000/000 000 000/000 000 010/000 000 000/000 000 000/000 000 000/000/
-T2	000 000 100/000 000 000/001 000/000 000 001/000 000 000 1--/000 000 000/ 001 000 000/000/
-T3	000 000/000 000 101/000 000/000 000 110/000 011 001 001/001 000/000 000 000/010 100/
-CT	000 000/000 000 101/001 000/000 000 111/000 011 001/001 000/001 000 000/ 010/
S3-T1	000 000 001/000 000/000 000/000 000 000 000/000 000 000 000/000 000/000 000 000/000 000 000/000 000 001/001 000 010/000 000 000/
-T2	000 000 000 000/000 000/000 000/000 000 000 000/000 000 000 000/000 000 100/ 100 000 000/000 000 000/000 000 100/010 001 000/000 000 000/
-T3	001 000 000/010 000 000/000 000 000/000 000 000/100 000 000/000 000 010/001 000/000 100 000/100 000 000/000 010 000/000 010 000/
-CT	001 000 001/010 000/000 000/000 000 000/100 000 000/000 000/101 000/000 100 000/100 000 101/011 011 010/000 010 000/
S4-T1	000 000 000/000 000/000 000 000/000 000/000 000 000/001 001 001/000 000 000/001 001 010/000 000 000/000 000 000/
-T2	111 111 111/001/000 000 000/010 000/000 000 000/000 100 100/000 000/000 000 000/001 010 000/000 000 000/
-T3	101 000 001/000 000/000 000 100 000/100 000/000 001 000/001 101 110/000 000 111/001 001/000 000 010 010/
-CT	111 111 111/001/000 000 100/110 000/000 001 000/001 101 111/000 000/001 001/001 010 010/111 111/
S5-T1	010 000 000/000 001 010/010 000/000 010 001/000 000 000 000/010 011/001 000 110/011 000/001 000 001/
-T2	001 000 000/001 000 011/111 001/100 010 001 00-/011 101 000 000/000 110/010 000 001/110 111 111/000 001 011/
-T3	000 110 001/010 000 011 00-/110 101/101 010 001/000 001 010 1--/011 001 000/ 000 111 010/010 000/010 001 000/
-CT	011 110 001/011 001 011/111 101/101 010 001/011 101 010/011 111/011 111 111/111 111/011 001 011/

Key: Error (1) Correct (0) Missing trial at end of block (-) Not recorded (?)
End of session (/)

Table B.8 Frequencies of Error Sequences Within Edited Blocks

		SEQUENCE (0=correct, 1=error)							
		000	001	010	011	100	101	110	111
Subject 1	Task 1	23	1	1	0	1	0	0	0
	Task 2	25	1	1	0	0	0	0	0
	Task 3	21	1	0	1	1	0	0	0
	Composite	16	3	2	0	1	0	0	1
Subject 2	Task 1	21	0	1	0	0	0	0	0
	Task 2	15	3	0	0	1	0	0	0
	Task 3	12	1	1	1	1	1	1	0
	Composite	8	3	1	1	0	1	0	1
Subject 3	Task 1	27	3	1	0	0	0	0	0
	Task 2	28	1	1	0	3	0	0	0
	Task 3	21	2	4	0	3	0	0	0
	Composite	15	2	2	1	3	2	0	0
Subject 4	Task 1	22	2	0	0	0	0	0	0
	Task 2	17	2	1	0	2	0	0	1
	Task 3	11	4	2	0	2	1	0	2
	Composite	6	5	1	0	1	0	1	2
Subject 5	Task 1	9	5	4	2	0	0	1	0
	Task 2	5	5	2	2	1	0	2	2
	Task 3	5	4	3	2	0	2	2	1
	Composite	0	2	0	6	0	1	0	2

References

Grant, D. A. Testing the Null Hypothesis and the Strategy and Tactics of Investigating Theoretical Models. Psychological Review, 1962, 69, 54-61.

NUREG/CR-1278, Handbook of Human Reliability Analysis with Emphasis on Nuclear Power Plant Applications, A. D. Swain & H. E. Guttmann, 1983.

NUREG/CR-2211, Modeling of Multiple Sequential Failures During Testing, Maintenance and Calibration, P. K. Samanta & S. P. Mitra, 1981.

Schurman, D. L. & Hawley, J. K. Conceptual Field Validation of Models of Multiple Sequential Failures During Testing, Maintenance and Calibration. ASA Report No. 709. Valencia, PA: Applied Science Associates, Inc., 1982.

APPENDIX C

MSF COMPUTER CODE AND USERS' GUIDE

This appendix contains the MSF computer code developed for the computation of the independent failure probability (p) and the dependence factor (k) using the moment estimation technique. The MSF computer code also evaluates the failure probabilities for different reliability configurations using p and k calculated from the input data. This appendix also provides the output of an example case run using the MSF computer code.

```
      PROGRAM MSF (INPUT,OUTPUT,TAPE1=INPUT,TAPE6=OUTPUT)
      DIMENSION F(2),FP(2,2)
     1,Z(2)
      COMMON NI,N,X1,X2,X3,X4,P12,P22,P13,P23,P33,P14,P24,P34,P44
      REAL K,NI,N
C READ CONSTANTS OF TH  PROBLEM
 1000 READ*,NCASE
      IF(NCASE.EQ.0) GO TO 300
      READ *,IFLAG
      READ*,NI,N,X1,X2,X3,X4,P,K
      WRITE(6,5) NCASE
    5 FORMAT(1H1,2X,*FINAL RESULTS FOR CASE*,I6)
      WRITE(6,6)
    6 FORMAT(1H ,2X,*-------------------------- --*
      WRITE(6,8)
    8 FORMAT(1H0)
      WRITE(6,4)
    4 FORMAT(1H0,2X,*INPUT DATA:*)
      PRINT *, NI,N,X1,X2,X3,X4,P,K
C
      IF(IFLAG.EQ.1) 20,30
20    CALL SKP(P,K)
      GO TO 300
30    CONTINUE
C  NOTE NI-IS THE NUMBER OF UNITS
      J=0
       XX=X1+X2+X3
       IF(XX.EQ.0.) GO TO 17
      K=(X2+X3+X4)/(X1+X2+X3)
        IF(K.GT.1.)K=1.
       GO TO 19
   17  K=1.
   19  CONTINUE
      P=1-(((N-(X1+X2+X3+X4))/N)**(1/N))
      WRITE(6,9)
    9 FORMAT(1H )
      PRINT*,P,K
C  THE ITERATION STARTS HERE
100   CONTINUE
      CALL FUN (F,P,K)
      WRITE(6,9)
      PRINT*,F
      J=J+1
      IF(ABS(F(1)).LT..01.AND.ABS(F(2)).LT..01) GOTO 200
      IF(J.LT.100) GOTO 120
      PRINT*,* SOLUTION WAS NOT REACHED IN 100 ITERATIONS-CHECK INPUT *
      PRINT*,* AND INITIAL SOLUTION*
      STOP
120   CONTINUE
      PP=P+0.00001
      PK=K+0.00001
      CALL FUN(Z,PP,K)
      FP(1,2)=-(F(1)-Z(1))*100000.
      FP(2,2)=-(F(2)-Z(2))*100000.
      CALL FUN(Z,P,PK)
      FP(1,1)=-(F(1)-Z(1))*100000.
      FP(2,1)=-(F(2)-Z(2))*100000.
```

```
      D=FP(2,2)*FP(1,1)-FP(2,1)*FP(1,2)
      IF(ABS(D).LT.0.000001)D=0.0001
      WRITE(6,111)D
 111  FORMAT(E10.2)
      K=K-(FP(2,2)*F(1)-FP(1,2)*F(2))/D
      P=P-(FP(1,1)*F(2)-FP(2,1)*F(1))/D
      IF(ABS(P).GE.1.)P=PP+0.001
      IF(ABS(K).GE.1.)K=PK+0.001
      IF(P.LT.0.)P=-P
      IF(K.LT.0.)K=-K
      WRITE(6,7)
      PRINT*,#   FOR THIS ITERATION P AND K ARE  #,P,K
      GOTO 100
 200  CONTINUE
      WRITE(6,7)
    7 FORMAT(1H0)
      PRINT*,#   AT THE SOLUTION  F(1) AND F(2) ARE  #,F(1),F(2)
      PRINT*,#   SOLUTION FOR P  AND  K  IS  #,P,K
      WRITE(6,10) NI,P,K
 10   FORMAT(1H0,2X,# SOLUTION FOR #,F2.0,# - UNIT SYSTEMS#,//,4X,#CALCU
     *LATED INDEPENDENT FAILURE PROBABILITY (P)=#,E10.4,/,4X,#CALCULATED
     * DEPENDENCE FACTOR (K)=#, F10.4)
       CALL SKP(P,K)
      GO TO 1000
  300  CONTINUE
       STOP
      END
```

```
        SYMBOLIC REFERENCE MAP (R=1)
ENTRY POINTS
  4141   MSF
VARIABLES      SN  TYPE            RELOCATION
  4605  D          REAL                              4606  F          REAL        ARRAY
  4610  FP         REAL        ARRAY                 4577  IFLAG      INTEGER
  4601  J          INTEGER                           4575  K          REAL
     1  N          REAL               / /            4576  NCASE      INTEGER
     0  NI         REAL               / /            4600  P          REAL
  4604  PK         REAL                              4603  PP         REAL
     6  P12        REAL               / /              10  P13        REAL              / /
    13  P14        REAL               / /               7  P22        REAL              / /
    11  P23        REAL               / /              14  P24        REAL              / /
    12  P33        REAL               / /              15  P34        REAL              / /
    16  P44        REAL               / /            4602  XX         REAL
     2  X1         REAL               / /               3  X2         REAL              / /
     4  X3         REAL               / /               5  X4         REAL              / /
  4614  Z          REAL        ARRAY
FILE NAMES        MODE
     0  INPUT      FREE        2054  OUTPUT   FREE          0  TAPE1          2054  TAPE6   FMT
```

```
      SUBROUTINE FUN(F,P,K)
      DIMENSION F(2)
      COMMON N1,N,X1,X2,X3,X4,P12,P22,P13,P23,P33,P14,P24,P34,P44
      REAL K,N1,N
      IF(N1-3.) 100,200,300
  100 P12=2*P*(1.-P)-K*P*(1.-P)
      P22=P*(1.-(1.-P)*(1.-K))
      F(1)=P12+2.*P22-(X1+2*X2)/N
      F(2)=P12+4.*P22-(X1+4.*X2)/N
      RETURN
  200 CONTINUE
      P13=2.*P*(1.-K)*(1.-P)**2+P*(1.-P)**2
      P23=P*(1.-P)-P*(1.-P)*(1.-2.*P)*(1.-K)
     1+P*(1.-P)*(1.-K)**2-P*((1.-P)**2)*(1.-K)**3
      P33=P-P*(1.-P)*(1.-K)-P*(1.-P)*(1.-K)**2+P*((1.-P)**2)*(1.-K)**3
      F(1)=P13+2.*P23+3.*P33-(X1+2.*X2+3.*X3)/N
      F(2)=P13+4.*P23+9.*P33-(X1+4.*X2+9.*X3)/N
      RETURN
  300 P14=P*(1.-P)**3+3.*P*(1.-K)*(1.-P)**3
      P24=P*(1.-P)**2+(3.*P-1)*P*(1.-K)*(1.-P)**2
     1+P*(2.+P)*((1.-P)*(1.-K))**2
     1-2.*P*((1.-P)*(1.-K))**3
      P34=P*(1.-P)+P*(1.-P)*(2.*P-1)*(1.-K)+
     1((1.-P)*P**2-(P*(1.-P))**2-P*(1.-P)**2)*(1.-K)**2
     1+(P*(1.-P)**3+P*(1.-P)-(P*(1.-P))**2)*(1.-K)**3
     1+P*((1.-P)**2)*(1.-K)**4-P*((1.-P)**2)*(1.-K)**5
     1+P*(1.-P)**3*(1.-K)**6
      P44=P-P*(1.-P)*(1.-K)-P*(1.-P)*(1.-K)**2
     1-(P**2)*(1.-P)*(1.-K)**3+P*((1.-P)**2)*(1.-K)**4
     1+P*((1.-P)**2)*(1.-K)**5-P*((1.-P)**3)*(1.-K)**6
      F(1)=P14+2.*P24+3.*P34+4.*P44-(X1+2.*X2+3.*X3+4.*X4)/N
      F(2)=P14+4.*P24+9.*P34+16.*P44-(X1+4.*X2+9.*X3+16.*X4)/N
      RETURN
      END
```

```
     SYMBOLIC REFERENCE MAP (R=1)
ENTRY POINTS
   3  FUN
VARIABLES      SN  TYPE          RELOCATION
    0  F            REAL      ARRAY   F.P.            0  K      REAL           F.P.
    1  N            REAL              / /             0  N1     REAL           / /
    0  P            REAL              F.P.            6  P12    REAL           / /
   10  P13          REAL              / /            13  P14    REAL           / /
    7  P22          REAL              / /            11  P23    REAL           / /
   14  P24          REAL              / /            12  P33    REAL           / /
   15  P34          REAL              / /            16  P44    REAL           / /
    2  X1           REAL              / /             3  X2     REAL           / /
    4  X3           REAL              / /             5  X4     REAL           / /
```

```
      SUBROUTINE SKP (P,K)
      COMMON NI
      REAL K,NI
      WRITE(6,7)
    7 FORMAT(1H0)
      WRITE(6,5)
5     FORMAT(1H0,2X,*CALCULATION OF DEPENDENT FAILURE PROBABILITY*)
C
      IF(NI-3.) 100,200,300
C
100   P12=(K*P) + ((P**2)*(1-K))
      WRITE(6,10) P12
10    FORMAT(1H0,2X,*1 OUT OF 2: G LOGIC SYSTEM =*,E10.4)
      RETURN
C
200   CONTINUE
      P13=P-P*((1-P)*(1-K))-P*((1-P)*(1-K)**2)+
     *P*((1-P)**2)*((1-K)**3)
      WRITE(6,20) P13
20    FORMAT(1H0,2X,*1 OUT OF 3: G LOGIC SYSTEM= *,E10.4)
C
      P23=P*(2-P) -((2*P)*(((1-P)**2)*(1-K)))
      WRITE(6,30) P23
30    FORMAT(1H0,2X,*2 OUT OF 3: G LOGIC SYSTEM =*,E10.4)
      RETURN
C
300   CONTINUE
      P14A=(1-(1-P))
      P14B=(1-(1-P)*(1-K))
      P14C=(1-(1-P)*((1-K)**2))
      P14D=(1-(1-P)*((1-K)**3))
       P14T=P14A*P14B*P14C*P14D
      WRITE(6,40) P14T
40    FORMAT(1H0,2X,*1 OUT OF 4: G LOGIC SYSTEM =*,E10.4)
C
      P24=P*(2-P)-(2*P)*(((1-P)**2)*(1-K))+P*((1-P)*
     *(P**2+P-2)*((1-K)**2))+(2*P)*(((1-P)**3)*
     *((1-K)**3))
      WRITE(6,50) P24
50    FORMAT(1H0,2X,*2 OUT OF 4: G LOGIC SYSTEM = *,E10.4)
C
      P34=(3-(3*P) + (P**2))*P-(3*P)*(((1-P)**3)*(1-K))
      WRITE(6,60) P34
60    FORMAT(1H0,2X,*3 OUT OF 4 : G LOGIC SYSTEM =*,E10.4)
C
400   CONTINUE
      RETURN
      END
```

SYMBOLIC REFERENCE MAP (R=1)

```
FINAL RESULTS FOR CASE      1
------------------------------


 INPUT DATA:
3. 132. 26. 5. 3. 0. 0. 0.
.09450918744005 .23529411176471
-.00RCC3944829257 -.04235407386593
 -.16E+C1

  FOR THIS ITERATION P AND K ARE  .09268781665552 .2984910575907
.000074359968020 41 .00068785199536 41

   AT THE SOLUTION  F(1) AND F(2) ARE  .000074359968020 41 .00068785199536 41
   SOLUTION FOR P  AND  K  IS  .09268781665552 .2984910575907

   SOLUTION FOR 3. - UNIT SYSTEMS
   CALCULATED INDEPENDENT FAILURE PROBABILITY (P)= .9269E-01
   CALCULATED DEPENDENCE FACTOR (K)=      .2985


  CALCULATION OF DEPENDENT FAILURE PROBABILITY

  1 OUT OF 3: G LOGIC SYSTEM=  .1847E-01

  2 OUT OF 3: G LOGIC SYSTEM = .8943E-01
```

NRC FORM 335 (6-83)

U.S. NUCLEAR REGULATORY COMMISSION

BIBLIOGRAPHIC DATA SHEET

1. REPORT NUMBER (Assigned by TIDC, add Vol. No., if any)	NUREG/CR-3837 BNL-NUREG-51786
2. Leave blank	
3. TITLE AND SUBTITLE	Multiple-Sequential Failure Model: Evaluation of and Procedures for Human Error Dependency
4. RECIPIENT'S ACCESSION NUMBER	
5. DATE REPORT COMPLETED	MONTH: March YEAR: 1985
6. AUTHOR(S)	P.K. Samanta, J.N. O'Brien, and H.W. Morrison
7. DATE REPORT ISSUED	MONTH: YEAR:
8. PERFORMING ORGANIZATION NAME AND MAILING ADDRESS (Include Zip Code)	Brookhaven National Laboratory, AUI Department of Nuclear Energy Engineering Technology Division Upton, NY 11973
9. PROJECT/TASK/WORK UNIT NUMBER	
10. FIN NUMBER	A-3219
11. SPONSORING ORGANIZATION NAME AND MAILING ADDRESS (Include Zip Code)	U.S. Nuclear Regulatory Commission Office of Nuclear Regulatory Research Washington, DC 20555
12a. TYPE OF REPORT	Formal
12b. PERIOD COVERED (Inclusive dates)	
13. SUPPLEMENTARY NOTES	

14. ABSTRACT (200 words or less)

This report provides an evaluation of the practicality, acceptability, and usefulness of using the Multiple Sequential Failure (MSF) model originally described in NUREG/CR-2111, 1981. The MSF model is described, discussed, and procedures for its use provided. The model was found to be practical, acceptable, and useful as a PRA tool for assessing the dependence due to human interactions with components in systems employing redundant components.

15a. KEY WORDS AND DOCUMENT ANALYSIS	15b. DESCRIPTORS
Probabilistic Risk Assessment	Calibration
Human Reliability Analysis	Human Errors
Sequential Failures	Human Factors
Maintenance	
Testing	

16. AVAILABILITY STATEMENT	17. SECURITY CLASSIFICATION (This report): Unclassified	18. NUMBER OF PAGES: 115
	19. SECURITY CLASSIFICATION (This page): Unclassified	20. PRICE: $