Previous Page

and human failures. Figure 5.11 is an event tree representation of operator actions involved in an offshore emergency shutdown scenario (Kirwan, 1990). This type of event tree is called an operator action event tree (OAET) because it specifically addresses the sequence of actions required by some initiating event. Each branch in the tree represents success (the upper branch) or failure (the lower branch) to achieve the required human actions described along the top of the diagram. The probability of each failure state to the right of the diagram is the product of the error and/or success probabilities at each node of branch that leads to the state. The overall probability of failure is given by summing the probabilities of all the failure states. The dotted lines indicate recovery paths from earlier failures.

In numerical terms, the probability of each failure state is given by the following expressions (where SP is the success probability and HEP the human error probability at each node):

F1 = [SP 1.1 + HEP 1.1 × SP 1.2] × SP 1.3 × SP 1.5 × SP 1.6 × SP 1.7 × HEP 1.8
F2 = [SP 1.1 + HEP 1.1 × SP 1.2] × SP 1.3 × SP 1.5 × SP 1.6 × HEP 1.7
F3 = [SP 1.1 + HEP 1.1 × SP 1.2] × SP 1.3 × SP 1.5 × HEP 1.6
F4 = [SP 1.1 + HEP 1.1 × SP 1.2] × SP 1.3 × HEP 1.5
F5 = [SP 1.1 + HEP 1.1 × SP 1.2] × HEP 1.3 × HEP 1.4
F6 = HEP 1.1 × HEP 1.2

Total failure probability T is given by

$$T = \mathrm{F1} + \mathrm{F2} + \mathrm{F3} + \mathrm{F4} + \mathrm{F5} + \mathrm{F6}$$

Further details about fault tree and event tree applications in quantitative risk assessment (QRA) are given in CCPS (1989b).

5.7. QUANTIFICATION

Because most research effort in the human reliability domain has focused on the quantification of error probabilities, a large number of techniques exist. However, a relatively small number of these techniques have actually been applied in practical risk assessments, and even fewer have been used in the CPI. For this reason, in this section only three techniques will be described in detail. More extensive reviews are available from other sources (e.g., Kirwan et al., 1988; Kirwan, 1990; Meister, 1984). Following a brief description of each technique, a case study will be provided to illustrate the application of the technique in practice. As emphasized in the early part of this chapter, quantification has to be preceded by a rigorous qualitative analysis in order to ensure that all errors with significant consequences are identified. If the qualitative analysis is incomplete, then quantification will be inaccurate. It is also important to be aware of the limitations of the accuracy of the data generally available

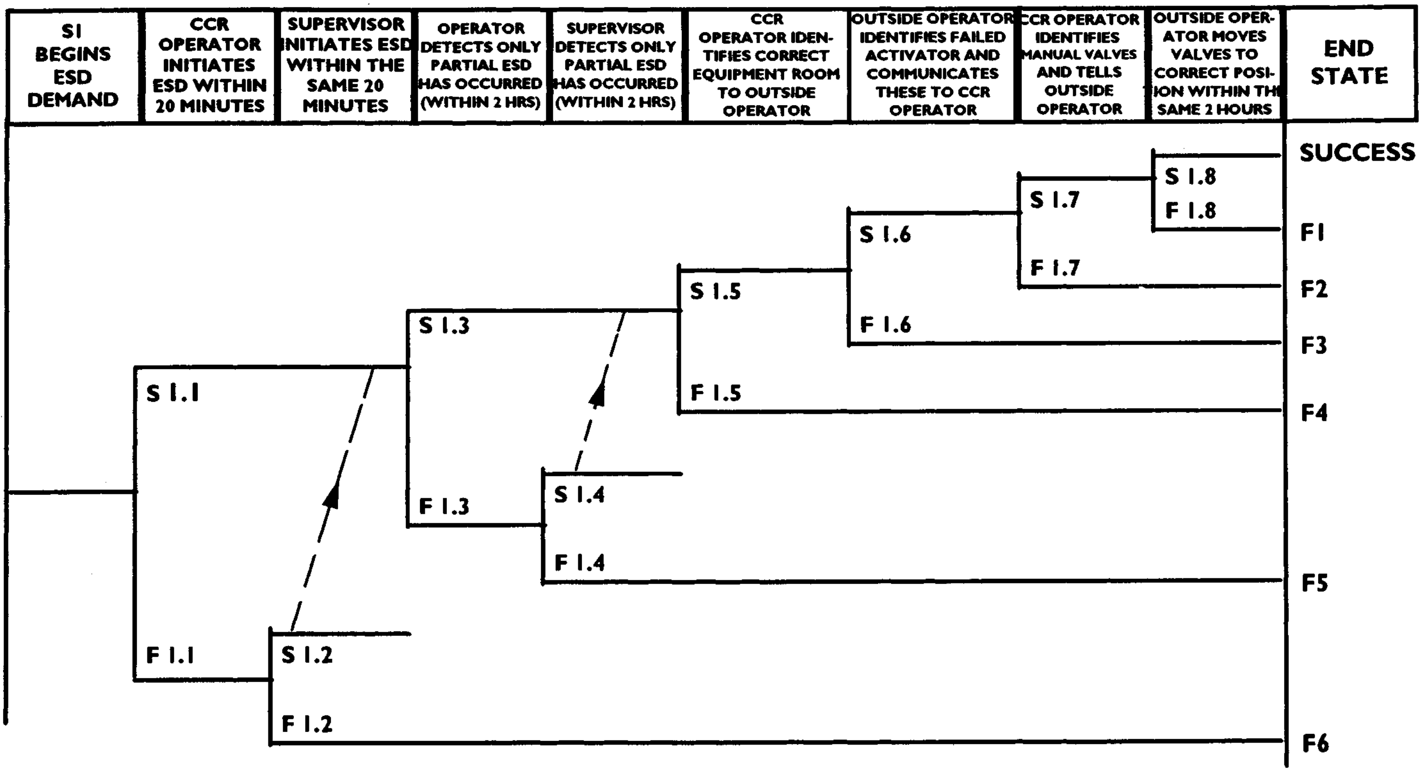


FIGURE 5.11. **Operator Action Tree for ESD Failure Scenario (Kirwan, 1990).**

for human reliability quantification. This issue is discussed in more detail in Chapter 6.

5.7.1. The Quantification Process

All quantification techniques follow the same four basic stages:

5.7.1.1. Modeling the Task

This involves analyzing the task of interest and identifying which aspects should be quantified. In some cases, the analyst will be interested in a probability for a discrete human action, for example, "what is the likelihood that the control room operator will close the feed supply valve within 30 seconds of an alarm?"

In other cases, the interest will be in quantifying a complete task, for example, "What is the probability that a lifeboat will be successfully launched?" In this case, quantification can be carried out at the global level of the whole task, or the task can be broken down to task elements, each of which is quantified (the decomposition approach). The overall probability of success or failure for the whole task is then derived by combining the individual task elements in some way.

Quantification at a global task level is essentially the same process as with a single discrete operation. A single probability is assigned without explicit reference to the internal structure of the task. There are arguments for and against both the global and the decomposition approach. The advantages of the decomposition approach are as follows:

- It can utilize any databases of task element probabilities that may be available.
- Recovery from errors in individual task steps can be modeled.
- Consequences to other systems arising from failures in individual task steps (e.g., the results of **alternative** actions as opposed to simply omitted actions) can be modeled and included in the assessment.
- Effects of dependencies among task steps can be modeled.

Advocates of the global approach would argue that human activities are essentially goal-directed (the cognitive view expressed in Chapter 2), and that this cannot be captured by a simple decomposition of a task into its elements. They also state that if an intention is correct (on the basis of an appropriate diagnosis of a situation), then errors of omission in skill-based actions are unlikely, because feedback will constantly provide a comparison between the expected and actual results of the task. From this perspective, the focus would be on the reliability of the cognitive rather than the action elements of the task.

On the whole, most quantification exercises have employed the decomposition approach, partly because most engineers are more comfortable with the analysis and synthesis approach, and partly because of the rather mechanistic model of human performance that has been the basis for most work in human reliability assessment.

5.7.1.2. Representing the Failure Model

The decomposition approach is used, it is necessary to represent the way in which the various task elements and other possible failures are combined to give the failure probability of the task as a whole. Generally, the most common form of representation is the event tree (see Section 5.7). This is the basis for THERP, which will be described in the next section. Fault trees are only used when discrete human error probabilities are combined with hardware failure probabilities in applications such as CPQRA (see Figure 5.2).

5.7.1.3. Deriving Error Probabilities for Task Steps

Error probabilities that are used in decomposition approaches are all derived in basically the same manner. Some explicit or implicit form of task classification is used to derive categories of tasks in the domain addressed by the technique. For example, typical THERP categories are selections of switches from control panels, walk-around inspections, responding to alarms and operating valves.

A basic error probability is then assigned to tasks in each category or subcategory. This probability may be derived from expert judgment or empirical data. It usually represents the error likelihood under "average" conditions. This probability is then modified by specifying a set of factors which tailor the baseline probability to the specific characteristics of the situation being assessed. Thus, a baseline probability of, say, 10^{-3} for the probability of correctly operating a valve under normal conditions may be degraded to 10^{-1} under the effects of high stress.

5.7.1.4. Combining Task Element Probabilities to Give Overall Task Failure Probabilities

During the final stage of the decomposition approach, the task element probabilities in the event tree are combined together using the rules described in Section 5.3.3 to give the overall task failure probability. At this stage, various corrections for dependencies among task elements may be applied.

5.7.2. Quantitative Techniques

To illustrate contrasting approaches to quantification, the following techniques will be described in detail in subsequent sections:

THERP	Techniques for human error rate prediction
SLIM	Success likelihood index method
IDA	Influence diagram approach

These techniques were chosen because they illustrate contrasting approaches to quantification.

5.7.2.1. Technique for Human Error Rate Prediction (THERP)

History and Technical Basis

This technique is the longest established of all the human reliability quantification methods. It was developed by Dr. A. D. Swain in the late 1960s, originally in the context of military applications. It was subsequently developed further in the nuclear power industry. A comprehensive description of the method and the database used in its application, is contained in Swain and Guttmann (1983). Further developments are described in Swain (1987). The THERP approach is probably the most widely applied quantification technique. This is due to the fact that it provides its own database and uses methods such as event trees which are readily familiar to the engineering risk analyst. The most extensive application of THERP has been in nuclear power, but it has also been used in the military, chemical processing, transport, and other industries.

The technical basis of the THERP technique is identical to the event tree methodology employed in CPQRA. The basic level of analysis in THERP is the task, which is made up of elementary steps such as closing valves, operating switches and checking. THERP predominantly addresses action errors in well structured tasks that can be broken down to the level of the data contained in the THERP Handbook (Swain and Guttmann, 1983). Cognitive errors such as misdiagnosis are evaluated by means of a time–reliability curve, which relates the time allowed for a diagnosis to the probability of misdiagnosis.

Stages in Applying the Technique

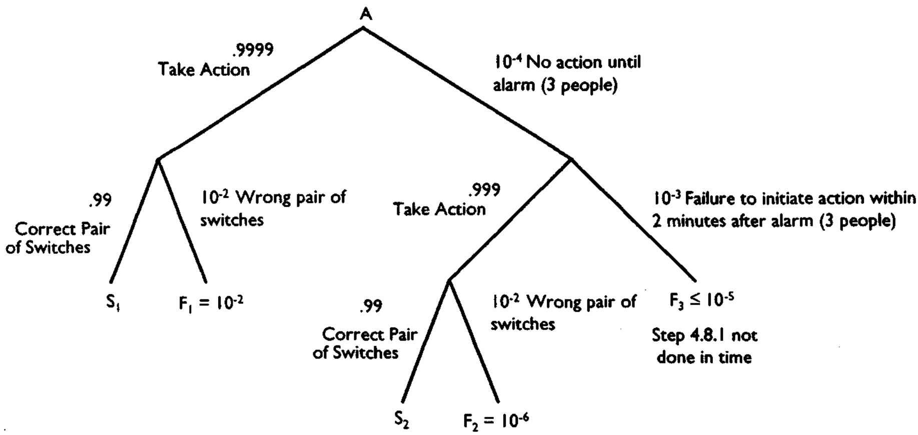
PROBLEM DEFINITION. This is achieved through plant visits and discussions with risk analysts. In the usual application of THERP, the scenarios of interest are defined by the hardware orientated risk analyst, who would specify critical tasks (such as performing emergency actions) in scenarios such as major fires or gas releases. Thus, the analysis is usually driven by the needs of the hardware assessment to consider specific human errors in predefined, potentially high-risk scenarios. This is in contrast to the qualitative error prediction methodology described in Section 5.5, where all interactions by the operator with critical systems are considered from the point of view of their risk potential.

QUALITATIVE ERROR PREDICTION. The first stage of quantitative prediction is a task analysis. THERP is usually applied at the level of specific tasks and the steps within these tasks. The form of task analysis used therefore focuses on the operations which would be the lowest level of a hierarchical task analysis such as that shown in Figure 5.6. The qualitative analysis is much less formalized than that described in Section 5.5. The main types of error considered are as follows:

- Errors of omission (omit step or entire task)
- Errors of commission
- Selection error
  —selects wrong control
  —mispositions control
  —issues wrong command
- Sequence error (action carried out in wrong order)
- Time error (too early / too late)
- Quantitative error (too little / too much)

The analyst also records opportunities to recover errors, and various performance shaping factors (called performance-influencing factors in this book) which will subsequently be needed as part of the quantification process.

REPRESENTATION. Having identified the errors that could occur in the execution of the task, these are then represented in the form of an event tree (Figure 5.12). This event tree is taken from Swain and Guttmann (1983). The branches of the tree to the left represent success, and to the right, failures. Although the event tree in Figure 5.12 is quite simple, complex tasks can generate very elaborate event trees. Error recovery is represented by a dotted line as in the event tree shown in Figure 5.11.



$$F_T = F_1 + F_2 + F_3 \approx 10^{-2}$$

FIGURE 5.12. **THERP Event Tree (Swain and Guttman, 1983).**

QUANTIFICATION. Quantification is carried out in the THERP event tree as follows:

- Define the errors in the event tree for which data are required. In Figure 5.12, these errors are:
  —No action taken until alarm (action omitted)
  —Failure to initiate action within 2 minutes of alarm
  —Wrong pair of switches chosen
- Select appropriate data tables in Swain and Guttmann (1983). This handbook contains a large number of tables giving error probabilities for operations commonly found in control rooms or plants, for example, selecting a switch from a number of similar switches. Because the handbook was originally written for the nuclear industry, the data reflect the types of operations frequently found in that industry. The source of these data is not defined in detail by the authors, although it appears to be partly based on the American Institute for Research human error database (Munger et al., 1962) together with plant data extrapolated and modified by the authors' experience.
- Modify the basic data according to guidelines provided in the handbook, to reflect differences in the assumed "nominal" conditions and the specific conditions for the task being evaluated. The major factor that is taken in to account is the level of stress perceived by the operator when performing the task.
- Modify the value obtained from the previous stage to reflect possible dependencies among error probabilities assigned to individual steps in the task being evaluated. A dependence model is provided which allows for levels of dependence from complete dependence to independence to be modeled. Dependence could occur if one error affected the probability of subsequent errors, for example if the total time available to perform the task was reduced.
- Combine the modified probabilities to give the overall error probabilities for the task. The combination rules for obtaining the overall error probabilities follow the same addition and multiplication processes as for standard event trees (see last section).

INTEGRATION WITH HARDWARE ANALYSIS. The error probabilities obtained from the quantification procedure are incorporated in the overall system fault trees and event trees.

ERROR REDUCTION STRATEGIES. If the error probability calculated by the above procedures leads to an unacceptable overall system failure probability, then the analyst will reexamine the event trees to determine if any PIFs can be modified or task structures changed to reduce the error probabilities to an acceptable level.

5.7.2.2. THERP Case Study

The case study that follows is reproduced with permission from the Chemical Manufacturers Association publication *Improving Human Performance in the Chemical Industry: A Manager's Guide,* Lorenzo (1990). Another CPI case study that uses THERP is documented in Banks and Wells (1992).

> Assume that the system described below exists in a process unit recently purchased by your company. As the manager, the safety of this unit is now your responsibility. You are concerned because your process hazard analysis team identified the potential for an operator error to result in a rupture of the propane condenser. You have commissioned a human reliability analysis (HRA) to estimate the likelihood of the condenser rupturing as the result of such an error and to identify ways to reduce the expected frequency of such ruptures

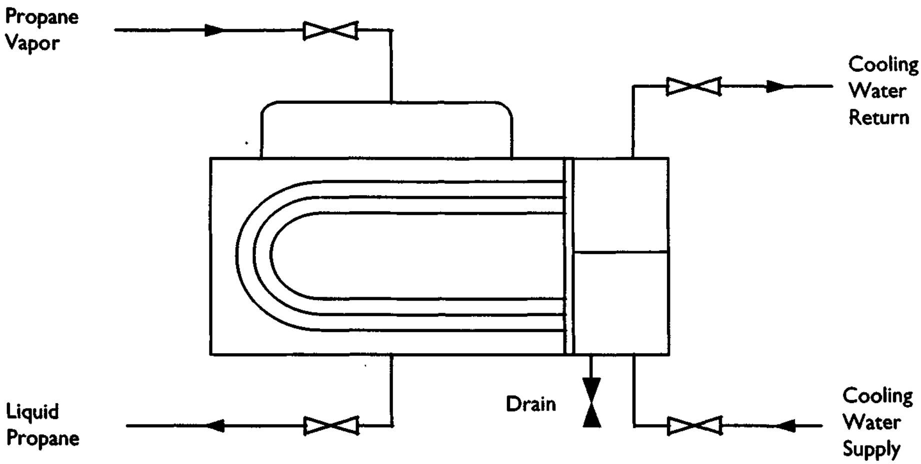


FIGURE 5.13. **Propane Condenser Schematic (Lorenzo, 1990).**

System Description

Four parallel propane condensers, one of which is illustrated in Figure 5.13, are designed with a 450-psig shell pressure rating and a 125-psig tube pressure rating. The propane vapor pressure is controlled at 400 psig; the cooling water flowing through the condenser tubes is normally maintained at 75 psig. Liquid propane flows out of the condenser as soon as it condenses; there is no significant inventory of liquid propane in the condenser. The two propane isolation valves for each condenser are rising-stem gate valves with no labels. The two water isolation valves for each condenser are butterfly valves with no labels. Their handwheel actuators have position indicators.

A tube has failed in one of the four condensers about once every three years. If a condenser tube fails, the affected condenser can be removed from service by closing four isolation valves (propane vapor inlet valve), liquid propane outlet valve, cooling water supply valve, and cooling water return valve). However, if a tube fails, it is essential that the operator close the two propane isolation valves before closing the two water isolation valves. Closing the two water valves first would allow pressure to build on the tube side of the condenser and rupture the tube head.

Analyzed System Conditions

- A tube has failed in the condenser.
- The low depropanizer pressure alarm has sounded in the control room.
- The experienced field operator has observed water and gas being emitted from the hydrocarbon vent at the cooling tower. The field operator shouts over the radio that a propane vapor cloud appears to be forming and moving towards the control room.
- The control room operator has directed the field operator to isolate the failed condenser as quickly as possible so that a unit shutdown will not be necessary.
- The operator must close the valves by hand. If a valve sticks, there is no time to go get tools to help close the valve—the process must be shut down.
- The field operator has correctly identified the condenser with the failed tube by the sound of the expanding propane and the visible condensation/frost on the shell.

Qualitative HRA Results

The first step of the analysis is to identify the human actions and equipment failures that can lead to the failure of interest. An HRA event tree (Figure 5.14) is then constructed to depict the potential human errors (represented by capital English letters) and the potential equipment failures (represented by capital Greek letters). The series of events that will lead to the failure of interest is identified by an F_1 at the end of the last branch of the event tree. All other outcomes are considered successes even though the propane release is not isolated in outcomes S_2 and S_3, so the process must be shut down.

Inspection of the HRA event tree reveals that the dominant human error is Error A: the operator failing to isolate the propane valves first. The other potential human errors are factors only if a propane isolation valve sticks open. Based on these qualitative results alone, a manager might decide to periodically train operators on the proper procedure for isolating a failed condenser and to ensure that operators are aware of the potential hazards. The manager might

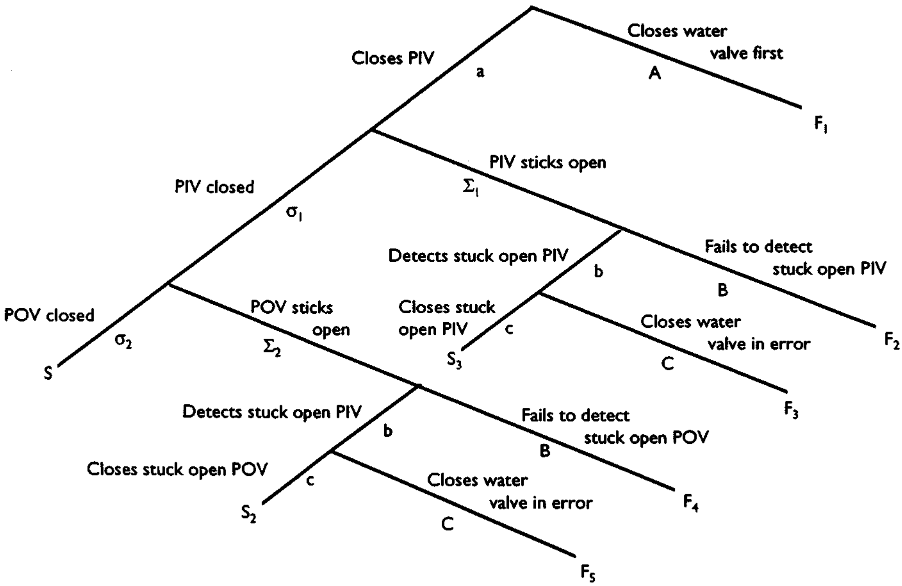


FIGURE 5.14. **HRA Event Tree for Improper Condenser Isolation (Lorenzo, 1990).**

also decide to require regular preventive maintenance on the propane isolation valves to help ensure that they will properly close when required.

Quantitative HRA Results

This manager requested quantitative results, so the analyst must estimate the probability of each failure or error included in the event tree. Data for all the failures and errors in this particular problem are available in tables in the *Handbook,* Swain and Guttman (1983). The analyst must modify these data as necessary to account for specific characteristics of the work situation, such as stress levels, equipment design features, and interoperator dependencies. Table 5.1 summarizes the data used in this problem.

There is a written procedure for condenser isolation, but it is normally a simple step-by-step task that is second nature to the operator and is performed from memory. However, under the threat of a potential vapor cloud explosion, the operator may forget to close the propane valves first (Error A). The HEP in Handbook *Table 20-7 #5 footnote (.01) is increased by a factor of 5 per* Handbook *Table 20-16 #6a to account for stress.*

The probability of a valve sticking open is unaffected by the operator's stress level, but the probability of the operator failing to detect the stuck valve (Error B) is increased. The HEP in *Handbook* Table 20-14 #3 is increased by a factor of 5 per *Handbook* Table 20-16 #6a.

TABLE 5.1
Events Included in the HRA Event Tree (Lorenzo, 1990)

FAILURE SYMBOL	FAILURE DESCRIPTION	ESTIMATED RROBABILITY	DATA SOURCE
A	Operator fails to close the propane valves first	.05	T20-7 #5 footnote × 5, per T20-16 #6a
Σ_1	Propane inlet valve sticks open	.001	T20-14 footnote
Σ_1	Propane outlet valve sticks open	.001	T20-14 footnote
B	Operator fails to detect a stuck valve	.025	T20-14 #3 × 5, per T20-16 #6a
C	Operator chooses to close the cooling water valves to stop the propane release	.25	T20-16 #7a

The third potential human error (Error C) is that the operator will decide to close the cooling water valves even though he/she diagnoses that a propane valve is not closed. The likelihood of such an error (a dynamic decision in a threatening situation) is listed as 0.25 in *Handbook* Table 20-16 #7a.

The analyst can then calculate the total probability of failure (F_T) by summing the probability of all failure paths (F_{1-5}). The probability of a specific path is calculated by multiplying the probabilities of each success and failure limb in that path. (*Note:* The probabilities of success and failure sum to 1.0 for each branch point. For example, the probability of Error B is 0.025 and the probability of Success b is 0.975.) Table 5.2 summarizes the calculations of the HRA results, which are normally rounded to one significant digit after the intermediate calculations are completed.

TABLE 5.2
Human Reliability Analysis Results (Lorenzo, 1990)

$F_1 = A$	$= 5.0 \times 10^{-2}$
$F_2 = a\Sigma 1B$	$= 2.4 \times 10^{-5}$
$F_3 = a\Sigma 1bC$	$= 2.3 \times 10^{-4}$
$F_4 = a\sigma 1\Sigma 2B$	$= 2.4 \times 10^{-5}$
$F_5 = a\sigma 1\Sigma 2bC$	$= 2.3 \times 10^{-4}$
$F_T = F_1 + \cdots + F_5$	$= .05$

Finally, the HRA analyst would calculate the expected frequency of condenser ruptures as a result of improper isolation. The frequency of condenser tube failures is 0.33 per year (1 every 3 years), and the calculated probability of improper isolation is 0.05. Multiplying these two numbers shows the expected frequency of improper isolation of a failed condenser is 0.017 per year, or about once every 60 years. The manager can use this number to help compare the costs and benefits of improvements proposed as a result of the HRA or other studies.

For example, the same process hazards review team that spurred the manager's original concern might have suggested (1) installing a pressure relief device on the tube side of the exchanger, or (2) removing the propane isolation valves (which would require that the unit be shut down in the event of a condenser tube failure). In addition, the HRA team may have suggested (3) increasing operator training and (4) more frequent maintenance of the propane isolation valves. Based on the quantitative HRA results and estimates of the consequences of a condenser rupture, the manager can decide whether the benefits of the proposed changes outweigh their costs. The manager can then choose the best way to apply loss prevention resources.

5.7.2.3. The Success Likelihood Index Method (SLIM)

History and Technical Basis

The SLIM technique is described in detail in Embrey et al. (1984) and Kirwan (1990). The technique was originally developed with the support of the U.S. Nuclear Regulatory Commission but, as with THERP, it has subsequently been used in the chemical, transport, and other industries. The technique is intended to be applied to tasks at any level of detail. Thus, in terms of the HTA in Figure 5.6, errors could be quantified at the level of whole tasks, subtasks, task steps of even individual errors associated with task steps. This flexibility makes it particularly useful in the context of task analysis methods such as HTA.

The basic premise of the SLIM technique is that the probability of error associated with a task, subtask, task step, or individual error is a function of the PIFs in the situation. As indicated in Chapter 3, an extremely large number of PIFs could potentially impact on the likelihood of error. Normally the PIFs that are considered in SLIM analyses are the direct influences on error such as levels of training, quality of procedures, distraction level, degree of feedback from the task, level of motivation, etc. However, in principle, there is no reason why higher level influences such as management policies should not also be incorporated in SLIM analyses.

In the SLIM procedure, tasks are numerically rated on the PIFs which influence the probability of error, and these ratings are combined for each task to give an index called the success likelihood index (SLI). This index is then converted to a probability by means of a general relationship between the SLI and error probability which is developed using tasks with known probabilities and SLIs. These are known as calibration tasks.

Stages in Applying the Technique

PROBLEM DEFINITION, QUALITATIVE ERROR PREDICTION AND REPRESENTATION. The recommended problem definition and qualitative error prediction approach for use with SLIM has been described in Section 5.3.1 and 5.3.2. The fact that PIFs are explicitly assessed as part of this approach to qualitative error prediction means that a large proportion of the data requirements for SLIM are already available prior to quantification. SLIM usually quantifies tasks at whatever level calibration data are available, that is, it does not need to perform quantification by combining together task element probabilities from a data base. SLIM can therefore be used for the global quantification of tasks. Task elements quantified by SLIM may also be combined together using event trees similar to those used in THERP.

QUANTIFICATION PROCEDURE. In order to illustrate the SLIM quantification method, the case study developed in the earlier part of the chapter based on the chlorine tanker filling example will be used. The following operations from Figure 5.6 will be used to illustrate the method.

2.1.3 Close test valve
4.1.3 Close tanker valve
4.4.2 Secure locking nuts
4.2.3 Secure blocking device on valves

- **Form groups of homogenous operations.**

The first stage is to group together operations that are likely to be influenced by the same PIFs. The four operations in the above set all involve physical actions for which there is no immediate feedback when incorrectly performed. Two of the operations, 4.1.3 and 4.4.2 are noted in Figure 5.8 as having significant consequences if they occur. It is legitimate to assume therefore, that the error probability will be determined by the same set of PIFs for all the operations in this set.

- **Decide on the relevant PIFs.**

Ideally, data bases will have been developed within a company such that predetermined PIFs are associated with particular categories of task. If this is not the case, the analyst decides on a suitable set of PIFs. In this example, it is assumed that the main PIFs which determine the likelihood of error are time stress, level of experience, level of distractions, and quality of procedures. (See Section 5.3.2.6.)

- **Rate each operation on each PIF.**

A numerical rating on a scale of 1 to 9 is made for each operation on each PIF. Normally the ends of the scale represent the best or worst PIF conditions.

For example, a high level of time stress would be represented by a rating of 9, which would imply an increased level of errors. However, in the case of level of experience, 9 would represent the optimal rating corresponding to a highly experienced operator. The fact that the same rating value can have a different significance with different PIFs needs to be taken into account by the analyst. With the computer program that is available for the SLIM technique, Embrey (1994), these adjustments are made automatically. The ratings shown in Table 5.3 are made for the operations.

These ratings can be interpreted as follows. In the case of the Time Stress PIF, all the operations have a high level of time stress, apart from **close test valve**, where stress is low. The operators are very experienced in carrying out all the tasks. Distractions are moderately high for **close test valve**, but otherwise low. Procedures are poor for **secure locking nuts** and **secure blocking device**, but above average for the other two tasks.

- **Assign weights if appropriate**

Based on the analyst's experience, or upon error theory, it is possible to assign weights to the various PIFs to represent the relative influence that each PIF has on all the tasks in the set being evaluated. In this example it is assumed that in general the level of experience has the least influence on these types of errors, and time stress the most influence. The relative effects of the different PIFs can be expressed by the following weights:

Time Stress	0.4
Distractions	0.3
Procedures	0.2
Experience	0.1

It should be noted that the analyst should only assign weights if he or she has real knowledge or evidence that the weights are appropriate. The assignment of weights is not mandatory in SLIM. If weights are not used, the technique assumes that all PIFs are of equal importance in contributing to the overall likelihood of success or failure.

TABLE 5.3
PIF Ratings

OPERATION	TIME STRESS	EXPERIENCE	DISTRACTIONS	PROCEDURES
Close test valve	4	8	7	6
Close tanker valve	8	8	5	6
Secure locking nuts	8	7	4	2
Secure blocking device	8	8	4	2

TABLE 5.4
Rescaled Ratings and SLIs

OPERATIONS	PIFs				SLIs
	TIME STRESS	EXPERIENCE	DISTRACTIONS	PROCEDURES	
Close test valve	0.63	0.88	0.25	0.63	0.54
Close tanker valve	0.13	0.88	0.50	0.63	0.41
Secure locking nuts	0.13	0.75	0.63	0.13	0.34
Secure blocking device	0.13	0.88	0.63	0.13	0.35
Weights	0.4	0.1	0.3	0.2	

- **Calculate the Success Likelihood Indices**

The SLI is given by the following expression:

$$SLIj = \sum R_{ij} W_i$$

where SLI_j is the SLI for task j; W_i is the normalized importance weight for the ith PIF (weights sum to 1); and R_{ij} is the rating of task on the ith PIF. The SLI for each task is the weighted sum of the ratings for each task on each PIF.

In order to calculate the SLIs, the data in Table 5.3 have to be rescaled to take into account the fact that the some of the ideal points are at different ends of the rating scales. Rescaling also converts the range of the ratings from 1 to 9 to 0 to 1. The following formula converts the original ratings to rescaled ratings:

$$RR = [1 - ABS\ (R - IP)]/[4 + ABS\ (5 - IP)]$$

where RR is the rescaled rating; R is the original rating, and IP is the ideal value for scale on which the rating is made.

The accuracy of this formula can be verified by substituting the values 1 and 9 for scales where the ideal point is either 1 or 9. The formula converts the original ratings to 0.0 or 1.0 as appropriate. Values of ratings between 1 and 9 are converted in the same way.

Using this formula on the ratings in Table 5.3 produces Table 5.4, which contains the rescaled ratings, the assigned weights for the PIFs and the calculated Success Likelihood Indices for each task.

- **Convert the Success Likelihood Indices to Probabilities**

The SLIs represent a measure of the likelihood that the operations will succeed or fail, relative to one another. In order to convert the SLI scale to a probability scale, it is necessary to calibrate it. If a reasonably large number of operations in the set being evaluated have known probabilities (for example, as a result of incident data having been collected over a long period of time), then it is possible to perform a regression analysis that will find the line of best fit between the SLI values and their corresponding error probabilities. The resulting regression equation can then be used to calculate the error probabilities for the other operations in the group by substituting the SLIs into the regression equation.

If, as is usually the case, there are insufficient data to allow the calculation of an empirical relationship between the SLIs and error probabilities, then a mathematical relationship has to be assumed. The usual form of the assumed relationship is log-linear, as shown below:

$$\log(\text{HEP}) = A\ \text{SLI} + B \tag{1}$$

where HEP is the human error probability and A and B are constants

This assumption is based partly on experimental evidence that shows a log-linear relationship between the evaluation of the factors affecting performance on maintenance tasks, and actual performance on the tasks, Pontecorvo (1965). In order to calculate the constants A and B in the equation, at least two tasks with known SLIs and error probabilities must be available in the set of tasks being evaluated.

In the example under discussion, it is found that there were few recorded instances of the test valve being left open. On the other hand, locking nuts are often found to be loose when the tanker returns to the depot. On the basis of this evidence and the frequency that these operations are performed, the following probabilities were assigned to these errors:

Probability of test valve left open = 1×10^{-4}
Probability of locking nuts not secured = 1×10^{-2}

These values, and the corresponding SLIs for these tasks (from Table 5.4), are substituted in the general equation (1). The resulting simultaneous equations can be used to calculate the constants A and B. These are substituted in the general equation (1) to produce the following calibration equation:

$$\log(\text{HEP}) = -2.303\ \text{SLI} + 3.166 \tag{2}$$

If the SLI values from Table 5.4 for the other two tasks in the set are substituted in this equation, the resulting error probabilities are as follows:

Task A: Probability of not opening tanker valve = 1.8×10^{-3}
Task B: Probability of not securing blocking device = 7.5×10^{-3}

- **Perform Sensitivity Analysis**

The nature of the SLIM technique renders it very suitable for "what if" analyses to investigate the effects of changing some of the PIF values on the resulting error probabilities. For example, there are high levels of time stress for both of the above tasks (rating of time stress = 8, best value = 1). The effects of reducing time stress to more moderate levels can be investigated by assigning a rating of 5 for each task. This changes the SLI, and if the new SLI value is substituted in equation (2) the probabilities change as follows:

Task A: Probability of not opening tanker valve = 5.6×10^{-5}
Task B: Probability of not securing blocking device = 2.4×10^{-4}

An alternative intervention would be to make the procedures ideal (rating = 9). Changing the ratings for procedures to this value for each task (instead of reducing time stress) produces the following results.

Task A: Probability of not closing tanker valve = 3.2×10^{-4}
Task B: Probability of not securing blocking device = 1.3×10^{-4}

Thus the effect of making the procedures ideal is an order of magnitude greater for Task B compared with Task A (see Table 5.5). This is because the procedures for Task A were already highly rated at 6, whereas there was room for improvement with Task B which was rated 2 (see Table 5.3).

TABLE 5.5
Effects of Improvements in Procedures on Error Probabilities Calculated Using SLIM

	ORIGINAL ERROR PROBABILITY	AFTER IMPROVEMENTS IN PROCEDURES	RATIO BEFORE/ AFTER IMPROVEMENTS
Task A	1.8×10^{-3}	3.2×10^{-4}	5.6
Task B	7.5×10^{-3}	1.3×10^{-4}	57

Conclusions

The SLIM technique is a highly flexible method that allows considerable freedom in performing what-if analyses. In common with most human reliability quantification techniques, it requires defensible data, preferably from a plant environment, to be effective. In the absence of such data, the calibration values have to be generated by expert judgments made by experienced plant personnel.

5.7.2.4. The Influence Diagram Approach

History and Technical Basis. The influence diagram approach (IDA) (also known as the sociotechnical approach to human reliability (STAHR) (see Phillips et al., 1990) is a technique that is used to evaluate human error probabilities as a function of the complex network of organizational and other influences that impact upon these probabilities. Unlike most other techniques, IDA is able to represent the effects of not only the direct influences of factors such as procedures, training, and equipment design on error likelihood but also the organizational influences and policy variables which affect these direct factors. As described in Phillips et al. (1990), it is possible to construct a generic Influence Diagram to represent these relationships. In the case study that will be used to illustrate the application of the influence diagram to human error probability evaluation, a more specific diagram (Figure 5.15) will be used, based on a study by Embrey (1992).

The basic steps in carrying out an IDA session are described in Phillips et al. (1990). A group of subject matter experts are assembled who have a detailed knowledge of the interactions between indirect and direct PIFs which determine error probability. The influence diagram is then constructed using insights from this expert group. Once the diagram has been developed, the experts are asked to assess the current state of the lowest level factors (i.e., project management and assignment of job roles in Figure 5.15). The assessment made is the probability (or "balance of evidence") that the factor being considered is positive or negative in its effects on error. This evaluation is performed on all the bottom level influences in the diagram, using scales similar to those used to evaluate PIFs described in Figure 3.1. Once these individual factors have been evaluated, based on an objective evaluation of the situation being assessed, the next stage is to evaluate the combined effects of the lowest level influences on higher level influences, as specified by the structure of the influence diagram.

This process is repeated for combinations of these variables on the factors that **directly** impact on the probability of success or failure for the scenario

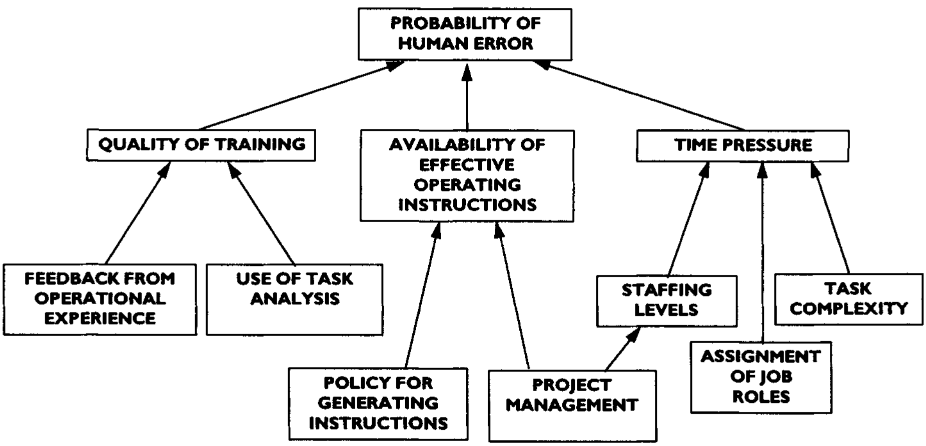


FIGURE 5.15 **Influence Diagram (Embrey, 1992).**

being evaluated. These numerical assessments are combined to give weights which are then used to modify unconditional probability estimates that the error will occur given various positive or negative combinations of the influences assessed earlier. The unconditional probability estimates have to be derived by another technique such as absolute probability judgment, SLIM, or from any field data that may be available.

Case Study

In the influence diagram for operator errors given in Figure 5.15, the main level 1 factors influencing the probability of error are quality of training, availability of effective operating instructions and time pressure on the operator. Two factors are specified as influencing the quality of training. These are the extent to which task analysis was employed to generate the training specification, and the use of feedback to modify the existing training regime in the light of operational experience. The availability of effective operating instructions is modeled as being dependent upon two policy factors. The first of these is the policy for developing instructions, which ensures that procedures are kept up to date, and are designed according to accepted standards. The other policy factor is project management, since this influences the early definition of work required, so that appropriate instructions will be available at the workplace when required.

Project management also influences the likelihood that staffing levels will be adequate for the tasks required. This latter factor, together with the extent to which appropriate jobs are assigned to individuals, and the complexity of the jobs, all influence the level of time pressure likely to be felt by the operator. The detailed calculations, which show how the probability of human error is influenced by changes in the sociotechnical factors in the situation, are given in Appendix 5A.

5.8. SUMMARY

This chapter has provided an overview of a recommended framework for the assessment of human error in chemical process risk assessments. The main emphasis has been on the importance of a systematic approach to the qualitative modeling of human error. This leads to the identification and possible reduction of the human sources of risk. This process is of considerable value in its own right, and does not necessarily have to be accompanied by the quantification of error probabilities.

Some examples of major quantification techniques have been provided, together with case studies illustrating their application. It must be recognized that quantification remains a difficult area, mainly because of the limitations of data, which will be discussed in Chapter 6. However, the availability of a systematic framework within which to perform the human reliability assessment means that despite data limitations, a comprehensive treatment of human reliability within CPQRA can still yield considerable benefits in identifying, assessing, and ultimately minimizing human sources of risk.

5.9. APPENDIX 5A: INFLUENCE DIAGRAM CALCULATIONS

Commentary on the Calculations

This commentary is provided to clarify the calculations in the following tables. In Table 1, the assessment team is asked to evaluate the evidence that feedback from operational experience is used to develop training. In order to make this evaluation, they will be provided with an "indicator" in the form of a scale specifying the nature of the evidence that should be taken into account. For example, the end of the scale defining the ideal situation would include conditions such as: "Results from operational experience fed directly to the training department," and "evidence that training regime is modified as a result of feedback." The other end of the scale would describe the worst case situation, for example, "No feedback from operational experience into training." In the example cited, the evidence strongly indicates that feedback is not used effectively in developing training.

1 What is the weight of evidence for feedback from operational experience in developing training?	
Good	Poor
.20	.80

2 What is the weight of evidence for use of task analysis in developing training?	
Used	Not Used
.20	.80

3	For Quality of Training				
If *feedback is:*	**and** *Task Analysis is:*	**then**	*weight of evidence that Quality of training is* *high is:*	*low is:*	Joint Weight (feedback x Task Analysis)
Good	Used		.95	.05	.04 (.20 x .20)
Good	Not Used		.80	.20	.16 (.20 x .80)
Poor	Used		.15	.85	.16 (.80 x .20)
Poor	Not Used		.10	.90	.64 (.80 x .80)
Unconditional Probability (weighted sum) that Quality of Training is high vs. low is:			.254	.746	

4 What is the weight of evidence that Policy for generating instructions is:	
Effective	Ineffective
.30	.70

5 What is the weight of evidence that Project Management is:	
Effective	Ineffective
.10	.90

6 For Availability of Effective Operating Instructions					
If *Policy for generating instructions is:*	**and** *Project Management is:*	**then**	*weight of evidence that operating instructions are* *available is:*	*not available is:*	Joint Weight (Policy x Project Management)
Effective	Effective		.90	.10	.03 (.30 x .10)
Effective	Ineffective		.60	.40	.27 (.30 x .90)
Ineffective	Effective		.50	.50	.07 (.70 x .10)
Ineffective	Ineffective		.05	.95	.63 (.70 x .90)
Unconditional Probability (weighted sum) that Effective Operating Instructions are available vs. not available is:			.255	.744	

Table 2 contains a similar assessment to Table 1 but for the use of task analysis. As illustrated in Table 3, the assessment team is then asked to evaluate the weight of evidence that the quality of training will be high (or low) given various combinations of the influencing factors feedback and use of task analysis. Of course, such evaluations are difficult to make. However, they utilize whatever expert knowledge is possessed by the evaluation team, and factor this into the analysis. They also allow the assessors to factor into their evaluations any interactions among factors. For example, the combined effects of poor feedback and nonuse of task analysis may degrade the quality of training more strongly than either influence in isolation. Each of the conditional assessments is then weighted by the results of stages 1 and 2 and the products added together to give an estimate of the unconditional probability that the training is adequate.

Similar assessments are performed to evaluate the probability that effective operating instructions are available (Table 6) that staffing levels are adequate (Table 9) and that time pressure will be high or low (Table 10). In this latter case, since three influences impact upon time pressure, eight joint assessments need to be made.

7 What is the weight of evidence for Assignment of Job Roles?	
Good	Poor
.50	.50

8 What is the weight of evidence for Task Complexity?	
High	Low
.60	.40

9 For Staffing Levels				
If *Project Management is:*	**then**	*weight of evidence that Staffing Levels are adequate is:*	*inadequate is:*	**Weight (Project Management) (from 5)**
Effective		.60	.40	.10
Ineffective		.20	.80	.90
Unconditional Probability (weighted sum) that Staffing Levels are adequate vs. inadequate is:		.24	.76	

10 For Time Pressure						
If *Staffing levels are:*	**and** *Assignment of Job Roles is:*	**and** *Project Management is:*	**then**	*weight of evidence for time pressure being high is:*	*low is:*	**Joint Weight (staffing levels x job roles x task complex.)**
Adequate	Good	High		.95	.05	.072 (.24 x .50 x .60)
Adequate	Good	Low		.30	.70	.048 (.24 x .50 x .40)
Adequate	Poor	High		.90	.10	.072 (.24 x .50 x .60)
Adequate	Poor	Low		.25	.75	.048 (.24 x .50 x .40)
Inadequate	Good	High		.50	.50	.023 (.76 x .50 x .60)
Inadequate	Good	Low		.20	.80	.015 (.76 x .50 x .40)
Inadequate	Poor	High		.40	.60	.023 (.76 x .50 x .60)
Inadequate	Poor	Low		.01	.99	.015 (.76 x .50 x .40)
Unconditional Probability (weighted sum) that Time Pressure is high vs. low is:				.3981	.6019	

Although these combined assessments are arduous, it should be noted that the evaluations of the effects of combinations of influences may be regarded as applicable across a range of systems, and hence would only need to be performed once for a generic model. The system specific evaluations would then be the simpler level 2 assessments set out in Tables 1, 2, 4, 5, 7, and 8. As discussed earlier, guidance for performing these assessments could be provided by the use of PIF scales delineating the conditions for the least and most favorable ends of the scales. Similar scales can be used to make direct evaluations of the level 1 influences, if the assessments described earlier are judged to be too difficult. Even if the full assessments are made, it is useful to compare these with the indirect assessments to check convergence.

The final stage of the procedure is to generate an overall unconditional probability of human error (Table 11). This is achieved by assigning probabilities of error to combinations of the three first level influences quality of training, availability of operating instructions and time pressure. These conditional probabilities are generic, in that they could apply to any system. They are made specific to the situation under consideration by multiplying them by the assessed probabilities of the level 1 influences, as derived from the earlier analyses. These products are then summed to give the overall unconditional probability of error occurrence in the situation being evaluated.

11 **For the task modeled**

If *Quality of Training is:*	**and** *Effective Operating Instructions are:*	**and** *Time Pressure is:*	**then**	*the probability of* *success is:*	*failure is:*	**Joint Probabilities (training quality x instructions x time pressure.)**
High	Available	Low		.99	.01	.0390 (.25 x .26 x .60)
High	Available	High		.978	.022	.0258 (.25 x .26 x .40)
High	N. available	Low		.954	.046	.1137 (.25 x .74 x .60)
High	N. available	High		.90	.10	.0752 (.25 x .74 x .40)
Low	Available	Low		.90	.10	.1145 (.75 x .26 x .40)
Low	Available	High		.78	.22	.076 (.75 x .26 x .40)
Low	N. available	Low		.54	.46	.3341 (.75 x .74 x .60)
Low	N. available	High		.00	1.00	.2209 (.75 x .74 x .40)
Assessed Unconditional Probability of success vs. failure is:				.58	.42	

The SLIM method described earlier is particularly suitable for the derivation of the conditional probabilities in Table 11, since it evaluates probabilities as a function of variations in PIFs that correspond to the level 1 factors used in this example. Each of the eight conditions in Table 11 can be treated as a separate task for evaluation by SLIM, using common weights for each factor across all conditions, but differing ratings to reflect the differing conditions in each case. SLIM requires calibration data to be supplied for the two end-point conditions, but this is considerably less onerous than evaluating probabilities for all conditions. Another source of probabilities to include in Table 11 would be laboratory experiments where the first level influencing factors were varied systematically.