QUANTIFICATION OF HUMAN ERROR RATES USING A SLIM-BASED APPROACH

Shan H. Chien
Andrew A. Dykes
John W. Stetkar
Dennis C. Bley
Pickard, Lowe and Garrick, Inc.
2260 University Drive
Newport Beach, CA 92660

Abstract

This paper presents an application of the success likelihood index methodology (SLIM) [1, 2, 3] for quantifying dynamic scenario-related human actions for use in a PRA. The application has been structured to make the assumptions, bases, and calculations leading to the quantitative evaluation of human error rates under plant transient conditions both scrutable and useful for use in risk management. It provides a structure in which assessment teams of operators and PRA analysts can provide feedback on the problems operators face and a means to prioritize corrective action. The utility of this procedure is expected to improve as the forms are updated to reflect the experience of previous applications and it is applied to a variety of situations.

Overview of SLIM

The success likelihood index methodology (SLIM) was developed under the sponsorship of Brookhaven National Laboratory and the U.S. Nuclear Regulatory Commission [1, 2, 3] to quantify operator actions in the plant response model of a probabilistic risk assessment. It is based on the assumption that the human error rate in a particular situation depends on the combined effects of a relatively small set of performance-shaping factors (PSF) that influence the operators' ability to perform the action successfully. PSFs account for both the plant conditions, or scenarios, under which the action must be performed and the psychological and cognitive state of the individuals performing the action. An example of a scenario-related PSF is the adequacy of time available to accomplish the action, while a psychological and cognitive PSF might address training and experience relative to the required action. The quantitative evaluation of the human error rate for the action is accomplished by judges who are assumed to be able to rank the PSFs in two ways:

- A numerical rating, r_i, of the degree to which the PSF helps or hinders the performance of the action.
- A ranking of the relative importance, or weight, w_i, of each PSF for influencing the reliability of the action.

An important assumption of SLIM is that the expert judges can select an appropriate set of PSFs and accomplish these two rankings independently of each other.

Once the ratings and weights have been obtained, a numerical success likelihood index (SLI) that represents the overall belief of the judges regarding the positive or negative effects of the PSFs on the likelihood of success for the action is calculated in accordance with the relation,

$$SLI = \sum_{i=1}^{I} w_i r_i \tag{1}$$

This numerical index is converted to a success rate for the action by assuming that it follows the relationship

$$\log_{10}(\text{success rate}) = a(SLI) + b \tag{2}$$

where a and b are calibration constants obtained by evaluating calibration tasks having "known" or "accepted" error rates in a similar manner. The basis and justification for the SLIM methodology are given in detail in References 1 through 3.

Summary of Application Features

This paper describes how the concepts of SLIM have been structured to facilitate the elicitation of expert opinion from plant operators and engineers for use in probabilistic risk assessments and risk management. The major features of this structure are:

- A set of seven predefined PSFs selected to span the spectrum of influences that might affect the operator's ability to accomplish the action.
- A set of forms to organize and document the information required to rate the action and its seven performance-shaping factors. These forms provide a qualitative assessment of the problems the operators may face while accomplishing the action.
- A rating scale that increases as the likelihood of failure increases. The ratings are then transformed into a failure likelihood index (FLI) in accordance with a relation that parallels Equation (1).

$$FLI = \sum_{i=1}^{I} w_i r_i \tag{3}$$

  Use of a larger rating for increasing failure likelihood permits a direct ranking of the contributors to human error. A large weight coupled with a high rating combines to make a large product, indicating a dominant contributor to the human error rate (HER).

  Once the FLI has been determined, it is converted into an HER using a formulation that parallels Equation (2).

$$\log_{10}(HER) = a(FLI) + b \tag{4}$$

- A Lotus 1-2-3 spreadsheet that can calculate the calibration constants a and b in Equation (2) and use them to determine the HER of actions from the ratings of the assessment team.
- A Lotus 1-2-3 spreadsheet that displays a ranking of contributors to the human error rate. This ranking is accomplished by multiplying the weight of the PSF by the numerical rating of the PSF by the assessment team.

Because the rating increases as the failure potential increases, the product of the weight and the rating becomes a direct measure of the relative contribution of that PSF to the human error rate of that action.

Implementation

Operator actions are selected for quantification by reviewing the plant event sequence diagrams and event trees to identify operator actions that impact plant risk. This process generally follows the methodology outlined in steps 1 and 2 of SHARP [4]. The definition of the operator action must consider the split fraction failure criteria for the scenario in which the action takes place.

A set of seven performance-shaping factors that address a spectrum of influences affecting operator actions has been defined. These seven PSFs cover most conditions the operator is expected to encounter. However, other PSFs may be used if warranted by the situation.

- **Plant Interface and Indications of Conditions.** This PSF relates the impact of the man-machine interface on the likelihood of success. It measures the degree to which the control room or the local conditions at the time when the action must be accomplished assist or hinder the operator in performing the action.
- **Significant Preceding and Concurrent Actions.** This PSF addresses the context of the modeled action. Preceding and concurrent actions can assist the action if they make it necessary and obvious to the operators. They can also divert the operators' attention from this action and cause a dependent failure. Lack of preceding actions may create a surprise effect that should be accounted for in this PSF.
- **Task Complexity.** This PSF rates the effect of multiple requirements on task success. It can range through the entire gamut of considerations to include coordination, multiple locations, remote operations, variety of tasks, and communications requirements. It also accounts for the availability of resources.
- **Procedural Guidance.** This PSF accounts for the extent to which plant procedures enhance the operator's ability to perform an action. The operator may have available not only step-by-step instructions, but also guidance on when the action has been correctly done.
- **Training and Experience.** This PSF measures the effect of the familiarity and confidence the operators have about the action. It accounts for the similarity of the action to previous operating transients. It also considers the frequency and depth of simulator and classroom training as it relates to this specific action.
- **Adequacy of Time to Accomplish Action.** This PSF considers the time required to complete the action compared with the time available and the effect on success. The rating reflects the confidence that the task can be accomplished in time to avert a change to a failed state. Depending on the definition of the action, the time required may include both the time required to diagnose the problem and the time to physically accomplish the action. The time available would then be measured from the first indication available to the operator.
- **Stress.** This PSF accounts for situations that may endanger the operator, damage or contaminate either the plant or the environment, or result in a long plant outage. Depending on its level, stress can serve as an incentive to accomplish the action, produce a reluctance to do it, or provide a diversion of attention that increases the likelihood of failure.

Each significant action is qualitatively evaluated on an operator response form designed to systematically lay out the context of the action, the cognitive tasks required to accomplish it, and those factors that influence the operator's ability to successfully accomplish it. This step is very similar to step 3 of SHARP [4]. Table 1 is a checklist that guides the completion of the form. It consists of three parts.

- Section A defines the action and establishes its context. It explicitly defines the tie between the plant risk model and the operator action being evaluated. The scenario up to the point at which the action is required provides the context. The split fraction failure criteria and time available define the plant state that will result from the operator's failure to accomplish the action and the estimated time available for the operator to act before the plant goes into that state. The time available is obtained from estimates of the rates of the physical processes involved.
- Section B breaks down the action into its cognitive elements. A cognitive element is a group of steps that can be completed before an operator must pause to obtain feedback from the plant or consider what to do next. This section explains how the action is accomplished with enough detail to identify potential problems. It is not necessary to list every step in the procedures the operators follow, but it is necessary to provide enough details, with references to procedures, to assist the evaluation team in identifying those performance-shaping factors that will influence the success or failure of the action.
- Section C summarizes how the PSFs influence the success or failure of the action. This section provides an opportunity to describe potential problems the operators might face while accomplishing the action. It can also delineate those things that can assist the operators.

The forms are most effective when they can be completed with active interaction with at least one senior reactor operator. The intent of the form is to accurately relate the problems the operators are expected to face if an accident scenario progresses to that point. The forms may be updated throughout the evaluation to reflect the insights of the evaluation teams. They improve the scrutability of the rating and can also provide suggestions for improving the procedures, training, or plant design.

Assessment teams consisting of operators and PRA team member quantitatively evaluate the actions. To make the process effective, the operators must understand that they are not being evaluated. Rather, the assessment is their opportunity to communicate the problems they face and provide suggestions for improving their ability to respond to the situation.

A set of descriptive scaling guides has been established to assist the rating of each PSF. An example of a scaling guide is given in Table 2. They are used as a reference to assist experts with different backgrounds in maintaining a consistent rating basis.

The descriptors on each scale are positioned to conform with the following general quantitative guidance for PSF ratings:

- A 0 corresponds to this PSF being "optimum" for assisting the operator team to accomplish the action in question.

Table 1. Guidelines for Completion of the Operator Response Form

HUMAN ACTION IDENTIFIER: ____________________ **TOP EVENT:** __________

HUMAN ACTION NAME: ______________________________

A. GENERAL DESCRIPTION

1. **Action Required**. Briefly state the action in general terms.
2. **Scenarios in Which Action Occurs**. State the broad context of the action. Identify initiating events and the previous response of the plant. Describe variations in the scenario that can affect the likelihood of success; e.g., system failures, previous operator errors, and conditions that could impact available time.
3. **Time Window Available**. State the physical or operational bases for time limitations the operator faces. Reference source of bases. Identify the plant change of state that indicates the end of available time.
4. **Split Fraction Failure Criteria**. Explicitly define the outcome of failing to accomplish the action correctly. Explicitly define boundary conditions to be evaluated. Objective is to succinctly summarize what is quantified.

B. TASK ELEMENTS

Provide sufficient detail to give a good picture of action, but it is not necessary to repeat every procedural step.

Task	Equipment	Location	Time Required	Comments

C. EVALUATION OF PERFORMANCE-SHAPING FACTORS

1. **Plant Interface and Indications of Condition**. Availability of alarms, instruments, and trend indications. Location of indications relative to required action. Quality of information: direct indication or interpretation required? Competing alarms and potential for confusion. Feedback to operator on correctness of response.
2. **Significant Preceding and Concurrent Actions**. Focus on required action or diverts attention? Is action expected or a surprise? Priority of action relative to other actions.
3. **Task Complexity**. Variety of subtask types and locations. Determine the level of cognitive process as skill, rule, or knowledge based. Number and qualifications of people required. Communications and coordination required. Potential demands on resources at time of action. Accessibility to the required plant equipment.
4. **Procedural Guidance**. Are they memorized or must they be read? How specific and applicable to this action? Assist in both diagnosis and response? Impact of EOPs on response. Existence of other supportive or conflicting guidance.
5. **Training and Experience**. Describe simulator training similar to action. Frequency of talk-throughs/walk-throughs on this action. Classroom or academic training. Similarity of training or experience to required action.
6. **Adequacy of Time To Accomplish Action**. Judge time available (Section A.3) relative to time required to complete (Section B, Time Required). Estimates by operators. Observations of simulator training.
7. **Stress**. Noise, vibration, radiation level, humidity, temperature, lighting, and other environmental stresses. Level of alertness at time of action (surprise factor). Perceived time available. Perceived threat or consequences. Toxic substance around working environment.
8. **Other**. List specific criteria.

- A 5 corresponds to conditions that neither significantly help nor hinder the performance of the action.
- A 10 corresponds to a condition when this PSF is hindering the performance of the action to the greatest extent possible.

The final group rating is obtained by consensus. Reasons considered in arriving at a consensus are recorded on the operator response form.

When the ratings of the actions have been completed, they are compared for consistency. Since human error rates will be calculated on the basis of these relative ratings, this review and update are essential.

The next step weighs the relative importance of the PSFs. The weight of a PSF relates the degree to which a change in the numerical rating of the PSF scale changes the operator's ability to accomplish the action. A PSF will have a large weight when a small change in the rating may produce a large change in the failure likelihood index. Conversely, if a large variation of the PSF rating scale has little impact on the likelihood of failure, the PSF will have little or no weight in determining the failure likelihood index. The relative weights of the PSFs affecting an action can be estimated by judging how much the rating of one PSF would be increased (made worse) to offset a decrease in the rating of another PSF by some convenient amount. Once the relative PSF weights are established, they are normalized to sum to 1.

When the weights have been established, the operator actions are classified into groups so that actions having similar PSF weights can be quantified together. The PSFs in different groups will have different normalized weights. However, within each group, only one set of normalized weights that is representative of the entire group will be used. This set can be obtained by averaging the weight of each PSF over the group or by reevaluating the PSFs considering the group as a whole. PSFs that are judged to have no significant influence on the likelihood of success of the group can be given a weight of zero.

Table 2. Performance-Shaping Factors and Scaling Guide

PSF: Significant Preceding and Concurrent Actions

Preceding and concurrent actions set the stage for the modeled action. They can assist the action if they make it necessary and obvious to the operators. They can also divert the operators' attention from this action or even cause failure. (If necessary, some strongly dependent failures may be accounted for by specific split fractions in the event trees.) Lack of preceding actions may create a surprise effect that should be accounted for in this PSF.

Scaling guidance:

— 0	Previous actions focus operators on the urgent need to act.
— 1	There are no distractions from this action; it is subject to close supervision and follow-up.
— 2	
— 3	Operators are alerted to the need for possible action and are expecting it.
— 4	Another step in standard or procedure-based responses.
— 5	Action is not a surprise, but previous actions create some competition for operator attention.
— 6	
— 7	This is one of many concurrent actions and could possibly be overlooked. Operator is taking recovery actions from one or two previous problems.
— 8	Operators are busy with other work or operators are in normal shift operations, and this is an unexpected, unusual transient.
— 9	Previous operator problems create an unusual situation.
— 10	The need to accomplish this action is unexpected and inconsistent with previous actions.

Consensus notes:

Examples of action groupings might be:

- Actions for which training and plant indications dominate, such as manual control of plant parameters.
- Actions for which time and preceding actions are most important, such as memorized immediate actions in response to a scram.
- Recovery actions for which the training and experience, the complexity of the action, and time available are important.

Grouping actions eliminates the need to quantify on the same scale actions that do not present the operators with the same types of problems. It permits the human action analyst to focus on those factors that most influence the error rate of the group of actions.

Calibration tasks are chosen for each group of actions. Calibration tasks have "known" or "accepted" values of HER and are influenced by PSFs with the same relative weights as the group of actions. The selection should include, if possible, one task that has a high likelihood of failure and one that has a low likelihood of failure. Calibration tasks are rated in the same way as the actions, and they form the basis for translating the FLI into human error rates. Calibration tasks are selected from human error rates determined from PRAs of other nuclear power plants using the results of human reliability experiments.

Additional calibration points can be obtained from best and worst case estimates of the influence of the PSFs on the group of actions being quantified. This technique involves estimating the likelihood of failure of a hypothetical action in which the group's PSFs combine to assist the operator (FLI = 0) or hinder the operator (FLI = 10) to the maximum extent possible. The result of this process is an estimate or the range of human error rates over which the group of actions may vary.

The failure likelihood index and HER of the actions are determined using a Lotus 1-2-3 spreadsheet that implements Equation (4). An example is given in Table 3. The failure likelihood index is calculated by multiplying the weight of each PSF by the rating of its contribution to the failure of the action and adding the products. The program obtains the constants for Equation (4) from a least squares fit of the calibration tasks and calculates the HERs for each group of actions using those constants.

The uncertainties of the HERs can be estimated from the uncertainties associated with the calibration tasks and the spread of HERs among several assessment teams. The result can be expressed in terms of the 5th, 50th, and 95th percentiles to represent the distribution of human error rates.

As a final check of overall consistency, the HERs of actions in each group should be compared with those of other groups. A judgment can then be made about whether differences in HERs are warranted by the differences in the scenarios and PSF ratings.

Example

The Lotus spreadsheets for a group of actions that primarily involve action to manually control reactor parameters during a transient are given in Tables 3 and 4. Table 3 illustrates the calculation of the human error rates.

Table 3. Example of the Spreadsheet to Quantify Human Error Rates Using a Failure Likelihood Index

PERFORMANCE-SHAPING FACTORS

	PLANT	ACTIONS	COMPLEXITY	PROCEDURES	TRAINING	TIME	STRESS	OTHER			
PSF Weights	0.3	0.1	0.1	0.15	0.2	0.1	0.05				
OPERATOR ACTIONS									FLI	HER	LOG(HER)
MAX HER	10	10	10	10	10	10	10		10.00	7.5E-01	-0.125
HEROF2	8	5	6	6	2	6	2		5.50	3.5E-02	-1.457
HERLI2	3	3	2	5	9	3	7		4.60	1.9E-02	-1.723
HEROL1	8	5	6	6	3	4	3		5.55	3.6E-02	-1.442
HEROF1	6	3	3	5	2	6	4		4.35	1.6E-02	-1.797
HEROL3	8	3	8	7	6	8	7		6.90	9.1E-02	-1.042
MIN HER	0	0	0	0	0	0	0		0.00	8.2E-04	-3.085
CAL 1	2	4	3	4	3	5	1		3.05	6.0E-03	-2.222
CAL 2	8	5	7	5	7	5	7		6.60	5.0E-02	-1.301
BEST CASE									0.00	1.0E-03	-3.000
WORST CASE									10.00	1.0E+00	-0.000

Regression Output:

Constant	-3.08525
Std Err of Y Est	0.162922
R Squared	0.989363
No. of Observations	4
Degrees of Freedom	2
X Coefficient(s)	0.2960667
Std Err of Coef.	0.0217072

Table 4. Example of the Spreadsheet to Quantify the Relative Contribution of Performance-Shaping Factors to Likelihood of Failure

PERFORMANCE-SHAPING FACTORS

	PLANT	ACTIONS	COMPLEXITY	PROCEDURES	TRAINING	TIME	STRESS	OTHER	TOTAL FLI
PSF Weights	0.3	0.1	0.1	0.15	0.2	0.1	0.05	0	
OPERATOR ACTIONS									
MAX HER	3.0	1.0	1.0	1.5	2.0	1.0	0.5	0.0	10.0
HEROF2	2.4	0.5	0.6	0.9	0.4	0.6	0.1	0.0	5.5
HERLI2	0.9	0.3	0.2	0.8	1.8	0.3	0.4	0.0	4.6
HEROL1	2.4	0.5	0.6	0.9	0.6	0.4	0.2	0.0	5.6
HEROF1	1.8	0.3	0.3	0.8	0.4	0.6	0.2	0.0	4.4
HEROL3	2.4	0.3	0.8	1.1	1.2	0.8	0.4	0.0	6.9
MIN HER	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

Note: Entries are the product of the PSF weight and the PSF rating.

The user enters the weights of the PSFs directly into the indicated row. The PSF ratings of each action and calibration task are entered in the appropriate column. The accepted values of HER for the calibration tasks are entered in the column marked HER. Once this is completed, the calibration constants can be calculated with the least squares linear regression package available through the DATA command of Lotus 1-2-3. Once this has been accomplished, the HERs for the actions are automatically calculated and converted to standard scientific format by the right side of Equation (4) and the antilog formula in the columns labeled LOG(HER) and HER, respectively. Best case and worst case HERs are also calculated to provide the user with information regarding the range of HERs that result from the calibration.

Table 4 illustrates the use of the individual products, $w_i r_i$, to determine the dominant contributors to the human error rate. For example, plant interface and indications of condition is the largest contributor to operator error for all but one of the actions addressed by this group. This is a result of the relatively large weight given to obtaining feedback from plant indications required for manual control. Risk management now involves determining the specific reasons for high ratings and obtaining expert opinions on the improvement in the rating that could result from specific modifications.

Conclusions

The application methodology has the following advantages:

- It provides an organized method of eliciting the estimates of expert judges who are most familiar with the problems of accomplishing the actions.
- It provides a mechanism by which human error rates can be estimated within the context of the scenarios in which they will be performed.
- The step-by-step documentation of the consensus process makes the estimates scrutable and provides feedback for improving operator training and procedures.
- The set of forms and instructions to explain and implement the procedure enables its consistent application on a long-term basis and provides the flexibility to update and add actions as additional insight into operator actions is gained.

In other words, the structured application of SLIM presented in this paper can both qualitatively and quantitatively represent the problems the operators face in the context of the scenario in which they must function.

References

1. Embrey, D. E., "The Use of Performance Shaping Factors and Quantified Expert Judgement in the Evaluation of Human Reliability: An Initial Appraisal," Brookhaven National Laboratory, prepared for the U.S. Nuclear Regulatory Commission, NUREG/CR-2986, May 1983.

2. Embrey, D. E., et al., "SLIM-MAUD: An Approach to Assessing Human Error Probabilities Using Structured Expert Judgment," Brookhaven National Laboratory, prepared for U.S. Nuclear Regulatory Commission, NUREG/CR-3518 (two volumes), March 1984.

3. Rosa, E. A., et al., "Application of SLIM-MAUD: A Test of an Interactive Computer-Based Method for Organizing Expert Assessment of Human Performance and Reliability," Brookhaven National Laboratory, prepared for U.S. Nuclear Regulatory Commission, NUREG/CR-4016, September 1985.

4. Hannaman, G. W., et al., "Systematic Human Action Reliability Procedure (SHARP)," NUS Corporation, prepared for Electric Power Research Institute, NP-3583, June 1984.