Reliability Engineering and System Safety **35** (1992) 127–136

INTENT: a method for estimating human error probabilities for decisionbased errors

David I. Gertman, Harold S. Blackman, Lon N. Haney, Karen S. Seidler & Heidi A. Hahn

Idaho National Engineering Laboratory, Idaho Falls, Idaho 83415, USA

(Received 22 August 1990; accepted 5 November 1990)

The development of a method, INTENT, for estimating probabilities associated with decisionbased errors is presented. These errors are not ordinarily incorporated into probabilistic risk assessments (PRAs) due to both the difficulty in postulating such errors and to the lack of a method for estimating their probabilities from existing data. By failing to include decisionbased errors in their analyses, most PRA practitioners seriously underestimate the true contribution of human actions to systems failure. This paper attempts to extend the identification of such errors and to quantify them. Two sources, Nuclear Computerized Library for Assessing Reactor Reliability (NUCLARR) and licensee event reports (LERs) were reviewed and two methods, HSYS and SNEAK, were used to identify a generic list of twenty potential errors which may be manifest as erroneous acts. Four categories of influence emerged from the data: consequence, attitudes, response set, and dependency. Corresponding human error probabilities (HEPs) for each error were generated by expert judgment methods. Lower and upper bounds for the HEPs for each error were determined by positing a situation reflecting optimized and degraded performance shaping factors, respectively. To allow analysts the opportunity to refine these extreme HEP values when evaluating a particular scenario of interest, normalization procedures were conducted and generic importance weights were computed for each of 11 performance shaping factors (PSFs) believed to affect the 20 decisionbased errors. It is believed by the authors that PSFs constitute a performance influence which, in some cases, such as in that for training, can serve to either augment or reduce the intellectual resources used by people to successfully accomplish tasks. These derived importance weights are used in conjunction with situation specific PSF ratings to compute a composite PSF score which, in turn, is mapped onto an HEP distribution. Distribution assumptions are presented and a function defining the relationship between composite PSF scores and HEPs is presented for use by the analyst.

1 INTRODUCTION

The Technique for Human Error Rate Prediction (THERP)[1] is widely used in the nuclear industry for modeling human performance and for providing failure rate estimates for both errors of omission and commission. The types of commission errors addressed by THERP are restricted to those commonly referred to as errors of selection and execution, which encompass actions performed incorrectly or at inappropriate times. The inadvertent selection of the wrong control from a bank of controls or the misreading of a display or indicator are representative of these types of errors. THERP, though, does not address an important subset of errors of commission known as errors of intention. The distinction between errors of commission and errors of intention seems to be acknowledged by a number of researchers. As Hollnagel *et al.*[2] suggest, design errors may be reduced by design reviews and failure analyses, erroneous actions can not be reduced by the same methods. They further contend that to some extent these types of errors are unavoidable. If this is the case, it behoves us to identify (1) the entry requirements necessary prior to the occurrence of errors of intention, and (2) the expected failure rate once entry requirements have been met.

Reliability Engineering and System Safety 0951-8320/91/$03·50

Errors of intention are therefore related to cognitive functions. The ability to reason, evaluate actions and estimate their consequences, as well as to weigh evidence, interpret rules, and regulations is part of man's make-up as a rational creature.[3] While errors related to these knowledge-intensive activities may result in inappropriate actions, they stem from erroneous decision-making, poor understanding of rules and procedures, and inadequate problem-solving. For instance, an incorrect control action resulting from the decision to operate outside of procedures or from the inappropriate application of a heuristic or misunderstanding of system relationships is considered to be an intention error. The emphasis here is that the error basis lies more in the thinking than in the doing.

Like errors of selection or execution, errors of intention may be active or latent, that is, their full impact may be manifested immediately or can lie dormant until triggered by some insidious combination of hardware and human actions.[4]

Previous research on attention suggests that the origins of cognitive error can be due to persons being either 'data limited' or 'resource-limited'.[5,6] Data limited refers to the quality of the data present, i.e. the signal to noise ratio, and may be linked to performance shaping factors (PSFs) such as procedures quality and the human–machine interface (HMI). The clarity of procedures can influence performance as can the degree of precision present in local instrumentation. Resource limitation, refers to people's capacity for processing information. Training may enhance people's ability to process and mentally file information for later use. Similarly, training may reduce the probability for a crew to misapply either procedures or heuristics. For example, highly trained crews would be less likely than poorly trained crews to solve the more minor of two faults. Rather, their attention may assumed to be focused, as it should, upon the consequence of each of the faults. An additional body of research suggests that the manner in which people perceive gains and losses associated with various outcomes also influences their behavior (decision making and actions).[7]

Although the use of computer simulation of human intention formation as a tool to assist the analyst in determining errors of intention is only emerging, at least one recent effort, the Cognitive Environment Simulation (CES), holds some promise for simulating the types of cognitively-based errors which may arise during accident scenarios.[8]

The cognitive nature of errors of intention and the fact that they can result from a wide range of factors, such as poor training or a poor safety culture at a facility, make it difficult to model and quantify them. So while there exists a large amount of anecdotal information on errors of intention, there is little corresponding probabilistic information that would support a human reliability analysis (HRA). Without such a quantitative data source, however, errors of intention will be underrepresented in HRA and PRA studies.

The present study was aimed at developing a method for the estimating probabilities associated with committing errors of intention. A list of errors of intention was generated and expert judgment was then used to assess the likelihood of these errors occurring under various conditions. These conditions were defined by different levels of PSFs. The generated list of errors may not be exhaustive, however, it provides a foundation on which to build a more complete database as field data becomes available. In the interim, it may be used by human reliability analysts to account for errors of intention.

The sections which follow present a brief review of several previous efforts to quantify errors of intention, followed by the development of INTENT, and then, by example, demonstrates how the analyst may apply INTENT.

1.1 Previous attempts at quantification

Swain and Guttman[1] noted the 'absence of models that estimate the reliability of cognitive processes in applied situations'. In an effort to address the problem, they provided some estimates of the probability of cognitive errors occurring during fault diagnosis of both a single and double abnormal event as a function of the amount of time elapsed since annunciation of the fault(s) occurred. As the authors themselves acknowledged, however, the THERP model is very limited in its usefulness for modeling intention errors because it does not provide probabilities for cognitive errors which may occur during normal plant evolutions, inservice testing, or maintenance activities.

1.2 Time based constraints

More recently, Yeh and Teng[9] have attempted to account for errors of intention. They utilized a time-based correlational approach for determining the human error probabilities associated with an anticipated transient without scram (ATWS). Specifically, they focused their analysis on a sequence where a crew fails to initiate boron to control reactor reactivity. The HEP was modeled as a function of two competing variables, critical time and action time. Critical time represents the system time available to the crew while action time is the average time required by the crew to make its response to the plant transient. The model relies on probability density function for Weibull distributions for the critical time and action time and utilizes the maximum entropy estimator referenced in Swain and Guttman.[1]

Yeh and Teng's model is limited in its application to modeling errors of intention because it only considers a single performance shaping factor—time constraints. Errors of intention are not necessarily time dependent. In fact, at the risk of committing heresy, we suggest other PSFs may be much more appropriate. For a thorough review of existing limitations of time reliability correlation approaches to HRA, the reader is referred to Ref. 10. At a minimum, an adequate model of errors of intention must account for a wider range of situational factors than time alone.

1.3 A heuristic for quantification

Research performed by Ujita[11] also made use of a time dependent approach to model operator performance. Multibranch trees similar to the event trees used in PRA were used to describe response to a large break loss of coolant accident (LOCA) in a boiling water reactor (BWR). Based upon a review of the data contained in Swain and Guttmann,[1] Ujita[11] developed a heuristic to derive probability estimates for errors of commission that could be applied to the limbs of the multibranch event trees. The heuristic employed estimated errors of commission to be roughly one-tenth of the probability of a general error of omission. According to Ujita, errors of intention are considered to be extraneous acts representing a special case of errors of commission. Ujita proposed a β factor which allows for an extrapolation of failure estimates to errors of intention. The value for β results from Ujita's judgment that failure rates for all intention errors are one-tenth of the probability of a general error of commission. Although this method can assure the user of common distribution assumptions between errors of commission and intention, it is highly unlikely that the myriad variety of erroneous acts which people can commit have but one general failure rate.

The need for a method for identifying and quantifying errors of intention and their corresponding failure rates has been highlighted by recent research.[1,9–11] The sections which follow present a description of the development of a method—INTENT used in the present study to identify and then quantify a dataset for 20 errors of intention.

2 METHOD

The first stage of the study involved compilation of potential errors of intention pertinent to tasks at nuclear power plants. The second stage required the determination of corresponding failure rates. Figure 1 presents the task flow sequence used to establish a methodology for estimating errors of intention. An explanation of the methods and sources follows.

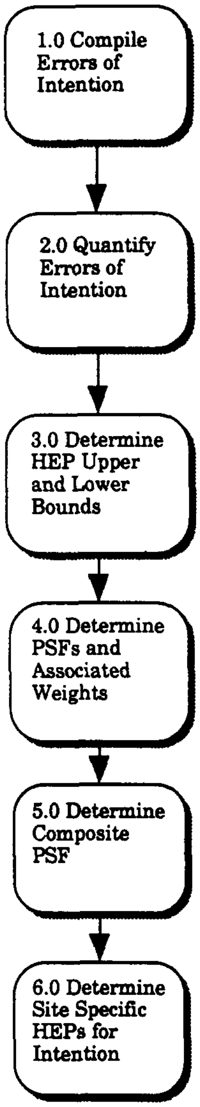

Fig. 1. Methodological flow for INTENT.

2.1 Compilation of errors

The identification of the errors of intention relied on applying SNEAK[12] and HSYS[13] to two potential sources of information, the Nuclear Computerized Library for Assessing Reactor Reliability (NUCLARR)[14,15] and licensee event reports (LERs). The search was conducted to determine the extent to which existing data for errors of intention might already be available in the open literature.

2.1.1 SNEAK analysis

SNEAK analysis is a methodology that was originally developed to identify pathways in electrical circuits that could lead to undesirable events. A modified approach was found to be useful for identifying errors of commission in nuclear power plant operator activities.[12] Essentially, the procedure involves developing flow diagrams or network trees to describe an activity of interest. A series of questions is then used to guide the analyst in the detection of potentially undesirable pathways, or 'sneak conditions', at each node of the tree. Different kinds of sneak conditions can occur at a node. There can be (1) *a sneak path,* where flow can take an unexpected path; (2) *sneak indication,* where misleading display output causes an inappropriate action; (3) *sneak label,* where a misleading control results in the inappropriate action; or (4) *sneak timing,* in which actions are inappropriately timed.

2.1.2 HSYS (Human-SYStem)

This is a method for investigating human performance in complex operational settings currently under development at the Idaho National Engineering Laboratory (INEL).[12] The structure is based upon the Input–Action model which describes human performance as preceeding through a series of five sequential stages, (1) input detection, (2) understanding of input meaning, (3) action selection, (4) action planning, and (5) action execution. A series of questions pertaining to the different factors that can affect performance at each of the five stages are organized hierarchically in tree and flowchart format. These formats guide users systematically in their examination and analysis of human performance in scenarios of interest. HSYS is currently being employed to assist in the examination and analysis of incidents at nuclear power plants.

2.1.3 NUCLARR

The first source examined for data related to errors decisionbased was NUCLARR. NUCLARR is a computerized data bank containing human error and hardware failure rate data relevant to the nuclear power industry. The NUCLARR system contains information about failure modes of human actions and equipment at three different levels of detail, (1) a systems level, (2) a components level, and (3) a displays and controls level. Instances of decisionbased were searched for at all three of these levels. Although errors of commission are represented at each level of the data bank in the NUCLARR data management system, no errors of commission specifically of the 'intention' type could be found. This largely reflects the fact that current HRA and PRA efforts do not report failure rates for decisionbased. SNEAK and HSYS were applied to the errors of commission resident in NUCLARR to see if these errors might also occur due to erroneous intention on the part of maintenance or operations personnel.

2.1.4 LERs

A second source of information, LERs, were also reviewed for instances where human error was the root cause of the event occurrence. Over 250 LERs involving 'human factors errors' were collected from the survey period 1985–1990 and reviewed. Of the identified root causes only those errors involving the execution of 'intentional' acts by personnel were extracted. That is, simple omissions or mistakes in executing a procedural step were omitted. The LER search was straightforward, in that the LERs contain descriptive, qualitative information. Item 3 in Table 1 whereby personnel violate procedure and rewire a breaker is an example of an item from the LER data source which would also be suggested by SNEAK analysis.

2.2 Categories of data

A number of the categories of sources for errors of intention given consideration in the present study were identified by applying SNEAK analysis to HSYS branchpoints to determine whether there was opportunity for an error of intention to occur. Individual errors were determined this way as well. For example, Figs. 2 and 3 presents a typical HSYS node reproduced from Ref. 13, the SNEAK questions which were applied to it, and the resulting error of intention and category which was identified. Specific error of intention categories identified include:

(1) *Action consequence.* The inclusion of consequence as part of the classification of errors reflects the consensus (the authors) that there is a relationship between consequence and the propensity for committing errors of intention. In some cases this relationship has been labeled a 'reluctance factor'.[16] Support for the influence of consequence on decision making is well researched. For a thorough review of the theoretical determinants of decision making and risk taking in groups in particular, the reader is refered to Whyte.[17] For example, the decisions made by crews which require selecting alternative courses of actions each of which may serve to mitigate an off-normal event may include the crew's ability to reach consensus regarding the risk of damaging one system in an effort to save another, more important (from a safety perspective) system. Such decisions go well beyond the simple stimulus–response model assumptions in use when describing errors of omission.

(2) *Crew response set, bias and interference.* Categorization of errors of intention by these sources reflects the influence that inhibition, response set and bias form as a function of experience, training, and previous learning. The effects of these influences on performance is well documented. The reader is refered to early studies by the Wurzburg school of Psychology where the

Table 1. Source categories and estimates of HEP upper and lower bounds for errors of intention

Source categories for errors of intention	HEP UB[a]	HEP LB[a]	EF
1.0 Action consequence			
1. Circumvent procedure with potentially catastrophic consequences, e.g. major ISLOCA with release.	7·5E-2	6·0E-5	35
2. Circumvent procedure with a minor consequence, e.g. a minor ISLOCA.	8·6E-2	3·3E-4	16
3. Tolerate an out of range situation with minor consequence.	3·6E-1	1·0E-2	6
4. Tolerate an out of range situation with moderate consequence.	1·5E-1	2·30E-3	8
2.0 Attitude leading to circumvention			
5. Violate procedure and reconfigure equipment.	8·3E-2	5·5E-4	12
6. Violate procedure and devise own formuli.	4·7E-2	1·6E-3	5
7. Checkers performing QA tolerate a discrepancy.	1·2E-1	1·2E-3	10
8. Common mode: failures due to poor safety culture.	2·0E-1	4·6E-3	7
3.0 Crew response set			
9. Misdiagnose given like symptoms. Capture sequence based on stimuli.	1·8E-1	1·3E-2	4
10. Right diagnosis–wrong response. Capture sequence based on response set.	2·2E-1	3·9E-3	8
11. Competing goal states leads to a wrong conclusion.	1·7E-1	8·9E-3	4
12. Symptoms noticed, but incorrect interpretation.	1·0E-1	4·2E-3	5
13. Correct actions taken during the wrong plant evolution.	3·2E-2	1·0E-3	6
14. Multiple fault situation, crew solves the more minor fault.	1·2E-1	1·2E-3	10
4.0 Resource dependencies			
15. Insufficient resources/instrumentation provided.	2·4E-1	7·4E-2	2
16. Crews consult inappropriate resources in emergency.	1·3E-1	1·9E-3	8
17. Inadequate communication results in improper actions.	2·0E-1	3·3E-3	8
18. Excessive task demands result in poor judgement.	2·9E-1	2·9E-2	10
19. Excessive task duration results in poor judgement.	9·0E-2	1·6E-2	2
20. Common mode: Poor judgements because procedures P&IDs, and operating conventions do not match.	2·9E-1	2·9E-2	3

[a] HEP UB refers to the case where the PSFs represent a worst case scenario; HEP LB refers to that situation where PSFs have been estimated to be optimal. Categories are not to be interpreted as errors themselves, but rather as sources for error. EF refers to error factors calculated for each of the 20 decisionbased errors.

task attitude of subjects (called the *Einstellung*) determined their perception for verbal recognition and abstraction tasks.

(3) *Attitudes leading to circumvention.* The inclusion of this category as a source brings to bear the fact that the manner in which individuals perceive the world (attitude) does have an influence on how they act in the world. Additionally, associations, perceptions and judgement may be changed by interests of the individual, as well as by the understanding the individual has of his immediate world. It suggests that perception, cognition, and actions associated with the two go well beyond simple stimulus–response paradigms.

(4) *Resource dependencies.* The category of *resource dependencies* is comprised of internal and external resources. Examples of the former include processing capacity, ability to withstand different types and degrees of stress and workload, and the limits of audition and vision. Examples of the latter include availability of emergency plans, operating procedures, instrumentation, and for highly complex systems perhaps computerized operator support systems as well.

2.3 Quantification considerations

In order to quantify the errors of intention identified by applying SNEAK analysis, and HSYS to the NUCLARR and the LER data sources, a Success Likelihood Index Methodology (SLIM)[18] session was first considered. SLIM produces an estimate of the HEP based upon group consensus of both relative PSF importance and PSF adequacy for a given scenario. The SLIM procedure, though, requires the inclusion of two calibration tasks with known failure rates and which share the same performance shaping factors as the other tasks in the group of errors of intention being considered. Additionally, SLIM requires the use of specific scenarios so that specific levels of salient PSFs can be defined and used to assess failure rates. In the present exercise, we considered a general class of errors identified through SNEAK analysis that were not plant or scenario specific, but which had specified, i.e. either optimal or severely degraded, PSF levels. Furthermore, it was impossible to identify tasks with known probabilities which were reasonable to use as calibration tasks for a SLIM session. Given the above constraints, the SLIM approach was discarded in favor of a direct estimation method which is presented below.

2.3.1 Determining HEP upper and lower bound estimates

First, seven human factors specialists trained in human reliability analysis rated the 20 errors of intention identified in the first stage of the study for their probability of occurrence under two extreme sets of conditions. The first condition required that participants considered a situation in which all conceivable performance shaping factors were optimized. The second condition dealt with the situation where all conceivable performance shaping factors were severely degraded. These two conditions were selected to assist in defining lower bound (LB) and upper bound (UB) HEP estimates, respectively. The PSF were not specified by the authors. Instead each analyst was asked to conceptualize performance in terms of the PSF that he or she felt would be most relevant to the errors of intention under consideration. A logarithmic probability scale was used to directly solicit these estimates. Analysts placed an 'x' on a probability scale for each error type corresponding to the HEP upper bound estimate and an '*' on the scale corresponding to the HEP LB. LBs and UBs were combined across analysts to get single pairs of estimates for each error type. The resulting bounds are presented as columns 2 and 3 of Table 1.

2.3.2 Determining PSFs and associated weights for each error type

The experts through group consensus techniques, identified and rated the influence of 11 PSFs on the failure rate for each of the 20 errors of intention. The 11 PSFs are: HMI: stress; skill; knowledge and rule based behavior (SRK); experience; safety culture; training; motivation; workload; supervision; communication; and procedures. Rating of PSF importance were generated independently by each analyst. For each error, the importance ratings were tabulated and then were normalized for each subject across the PSFs. Normalization was achieved in the following manner. For each error type, each expert's importance ratings were divided by the sum of those ratings for the error type. Next, for each error type and PSF, the resulting weights were then averaged across experts. This was done in order to account for individual differences in scale usage.

3 APPLICATION OF THE METHOD

Once the 20 errors were identified and PSF importance weights for each error determined, a method was needed whereby an HEP could be derived which would reflect the positive or negative influence of the PSFs on system performance. The most straightforward method for this was to allow for analysts to qualitatively evaluate, i.e. rate on a site and scenario specific basis on a favorableness continuum, the favorableness or unfavorableness of each PSF. For example, poorly administered or technically inadequate training would receive a highly unfavorable rating, i.e. a '1'. Well disciplined, comprehensive training, which made good use of on the job and classroom skill development would receive a highly favorable rating of '5'.

3.1 Determining site specific composite PSF ratings

The PSF importance weights are used in conjunction with the HRA analyst's ratings for the 11 PSFs. Ratings are generated with a site specific scenario in mind and used to determine a 'composite PSF' for an error type. The weights are error specific and ratings are site specific, e.g. ratings are specific to the nuclear setting under evaluation. The composite PSF for an error type is formed by multiplying each of the 11 PSF ratings by its corresponding weight, and then summing the results. Since the ratings are on a 1–5 scale, each composite PSF lies between 100 and 500. The use of the resulting composite PSF to determine the HEP estimate for errors of intention is described below.

3.2 Determining site specific HEPs for errors of intention

In general, the desired site-specific HEP (HEP_{is}) lies between the bounds for the error type presented in Table 1. For each error type, the composite PSF rating (F_{is}) lies between the values 100 and 500. (The subscript i refers to the error type and s to the specific site). By defining a mapping between the PSF rating scale and the HEP UB and LB, one can obtain a site-specific HEP. The mapping is defined such that a high PSF rating corresponds to a low HEP and vice versa. One way to define such a mapping is to consider the probability distributions of the composite PSF and the HEPs, and require that

$$P(F_{is} \leq x) = P(HEP_{is} \geq y) \qquad (1)$$

for any composite PSF, x, and corresponding site-specific HEP for intention, y. To completely define this mapping, the following additional assumptions are made:

1. The composite PSF, over various sites, has a uniform distribution. Thus, its 5th and 95th percentiles are 120 and 480, respectively.
2. The probabilities for errors of intention are log-normally distributed. The upper and lower bounds specified above are taken to be 5th and 95th percentiles for these distributions.

Table 2. Mean normalized PSF weights for twenty decisionbased errors

Error of intention no.	HMI	Stress	SRK	Experience	Safety culture	Training	Motivation	Work-load	Supervision	Communication	Procedures
1	6	8	9	9	11	12	7	8	11	7	11
2	7	8	9	9	10	10	7	9	11	8	11
3	9	8	9	9	10	11	8	8	10	7	11
4	8	9	10	9	9	11	9	8	10	5	12
5	9	8	9	10	10	11	7	9	11	7	9
6	9	8	10	10	11	10	8	9	11	7	9
7	8	9	10	9	10	10	9	10	11	6	8
8	11	10	11	11	5	11	8	11	7	8	9
9	11	11	9	10	6	10	7	11	9	8	10
10	6	12	11	10	9	11	9	10	9	6	8
11	11	10	11	10	6	11	8	10	8	7	8
12	9	9	9	9	7	10	8	10	9	9	9
13	8	10	10	10	7	11	7	9	9	11	8
14	14	7	9	9	9	9	7	10	11	9	6
15	10	11	9	9	8	10	7	10	9	9	8
16	9	9	8	10	9	11	8	9	8	13	8
17	8	12	8	10	8	10	7	13	8	9	7
18	9	11	8	10	8	10	9	11	9	7	7
19	11	8	10	10	9	11	6	7	8	8	13
20	7	6	7	8	23	10	6	7	9	8	8

If analysts rate each of the 11 PSFs on a five point Likert scale, as the authors suggest, then the exact 5th and 95th percentile values can be used in determining the site specific HEPs for errors of intention. Based on the uniform distribution assumed for composite PSFs, the left side of the equality presented above in eqn (1), simplifies to

$$P(F_{is} \leq x) = x - 100/400 \qquad (2)$$

Users should note that changes in the rating scale range would require a slight modification to this equation. Additionally, it is assumed that users will employ the sets of PSF importance weights presented in Table 2.

The evaluation of the right hand side of eqn (1) is based on assumption (2). Since the HEPs are lognormally distributed, their logarithms ($\ln y$) follow a normal distribution. The 5th and 95th percentiles of a normal distribution are its mean plus or minus 1·654 times its standard deviation, respectively. A similar statement applies to the 5th percentile. Thus, the mean (m) and standard deviation (sd) of the underlying normal distribution can be found by equating these percentiles with the ln UB and ln LB, respectively (from Table 2). As a function of m and sd, the right hand side of eqn (1) is

$$1 - \phi[(\ln(y) - m)/sd] \qquad (3)$$

where ϕ is the standard normal cumulative distribution function.

Solving for m and sd, equating the left and right sides of eqn (1), and then solving for y as a function of x produces the following equation:

$$y = \exp\{(\ln \mathrm{UB} - \ln \mathrm{LB}) \times \phi^{-1}[(500 - x)/400]/3{\cdot}29 + (\ln \mathrm{UB} + \ln \mathrm{LB})/2\} \qquad (4)$$

Here 'ϕ^{-1}' is the inverse of the standard normal cumulative distribution function. That is, it is the value a standard normal variable is less than or equal to, with probability $(500 - x/400)$. To use this equation, we substitute the error type and site-specific composite PSF rating for x and solve for y. The resultant y is the desired site-specific HEP.

4 FINDINGS

The 20 errors of intention identified by applying HSYS and SNEAK analyses to NUCLARR and LERs are presented in Table 1. The errors are grouped according to the following categories: action consequence; attitudes leading to circumvention; crew response set, bias, and interference; and resource dependencies. Definitions of these categories are provided in the Methods section. All four categories are thought to either distort or influence perception with the result that erroneous intention leading to erroneous action is a potential consequence.

Table 2 contains the errors of intention and their associated upper and lower bound HEP estimates as determined by expert judgment methods. Values range from $1{\cdot}0E-1$ to $6{\cdot}0E-5$ and represent failure rates for best and worst case PSF situations across a broad spectrum of plants. As depicted in Fig. 2,

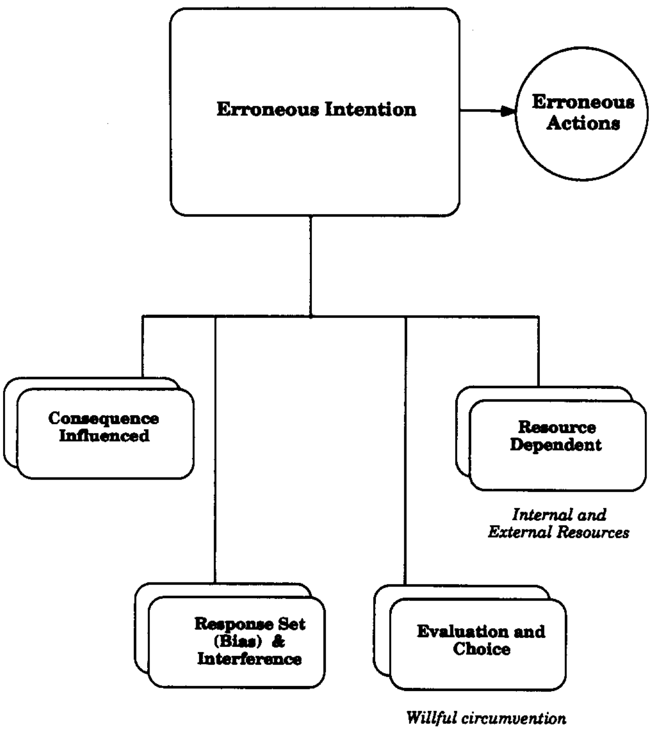

Fig. 2. Influences upon errors of intention as determined by HSYS and SNEAK analysis.

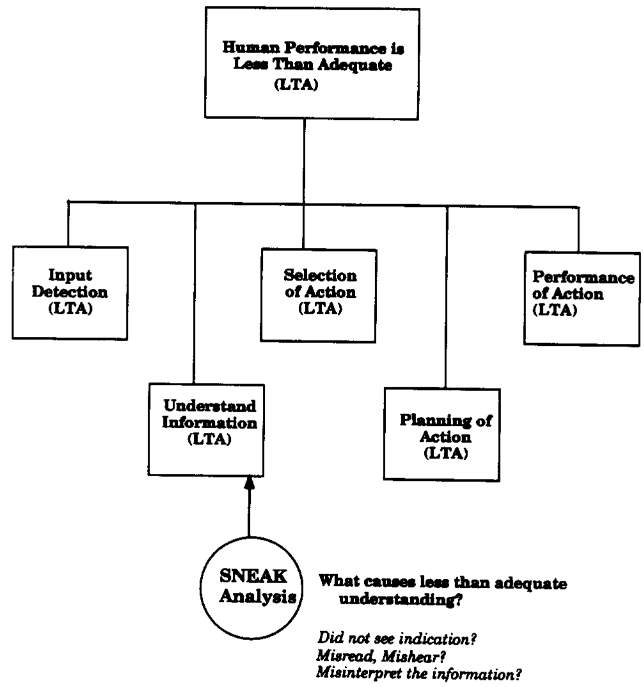

Fig. 3. SNEAK analysis application to HSYS for error identification. Note: Misinterpreting of information, revealed by SNEAK might be caused by a variety of factors. For example, response set could cause information to be misconstrued. The propensity for certain situations to arise more than others could condition personnel to misinterpret incoming information. This type of error appears in Table 2, Category 3, Item 9 as capture sequence based on stimuli.

erroneous thought or intention leads to poor evaluation and choice which, in turn, is manifest in erroneous action. Figure 3 presents SNEAK analysis application to an HSYS structure for purposes of error identification. Placing error likely scenarios in an HSYS structure was found to facilitate the SNEAK analysis.

The average error rate for the UB for error rates determined by the INTENT method is on the order of $1{\cdot}0E-1$, and the corresponding rate for the LB is $1{\cdot}0E-2$. Examining THERP data tables for comparative rates, i.e. those in the knowledge-based realm, shows errors on the order of $1{\cdot}0E-2$. The rates in both instances are high. However, the requirements for entry into a situation disposed to an error of intention is *restricted to instances discovered through SNEAK analysis,* and therefore the observed instances for such occurrences is infrequent. Once personnel find themselves in a tenuous situation the rates proposed are, we feel, realistic. For example, most of the time crews who have been trained to operate to procedures which have been validated can be expected to perform in an admirable fashion. However, given an off-normal situation whereby crews are forced to design a metaprocedure or perform back of the envelope calculations on the fly, the rates suggested in Table 2 are what one would expect to find.

Each of the errors should be quantified individually by the analyst and resulting HEPs are meant to serve as input to either fault trees or HRA event trees. In multiple fault instances, it is suggested that the analyst identify a dominant cognitive error appropriate to the crew's response to each individual fault. In situations where 'or' rather than 'and' gate logic is used, it should be noted that rates on the order of $4{\cdot}0E-1$ may be observed. We believe this to be a realistic failure rate for use in risk calculations for certain infrequent situations. For example, those situations for which SNEAK analysis has determined that the crew is unprepared, and the combination of events unprecedented and untrained for, the proposed failure rates may be less than conservative.

Table 2 presents the 20 sets of 11 normalized PSF weights for each of the errors of intention developed as part of this study. The weights range from 5 to 13 and are selected as a group dependent upon the error of intention identified for review by the analyst. An example is presented below to describe how the data tables may be used to secure an HEP value for an error of intention.

5 USING THE INTENT DATA TABLES: AN EXAMPLE

Table 3 presents the ratings from a hypothetical example where an HRA analyst has made use of the

Table 3. Ratings and calculations for procedural violations (error no. 4) at plant 'X'

PSF	PSF rating	PSF weights	Rate × weight
HMI	2	8	16
Stress	3	9	27
S/R/K	2	10	20
Experience	3	9	27
Safety Culture	2	9	18
Training	2	11	22
Motivation	4	9	36
Workload	4	8	32
Supervision	2	10	20
Communication	3	5	15
Procedures	2	12	24
			(Total R × W = 257)

Table 4. Estimating the HEP for procedural violation (error no. 4) at plant 'X'

From Table 2: UB = $4{\cdot}7E-2$, LB = $1{\cdot}6E-3$, lnUB = $-3{\cdot}057$, lnLB = $-6{\cdot}4378$

Eqn 4:

$$\mathrm{HEP}_{is} = \exp\{(\ln\mathrm{UB} - \ln\mathrm{LB}/3{\cdot}29)\quad \phi^{-1}(500 - x/400)) + (\ln\mathrm{UB} + \ln\mathrm{LB}/2)\}$$

$$\exp\{((-3{\cdot}0576) - (-6{\cdot}4378)/3{\cdot}29) \times \phi^{-1}(500 - 257/400) + (-3{\cdot}0576 + 6{\cdot}4378)/2\}$$

$$= \exp\{(1{\cdot}0274)\phi^{-1}(0{\cdot}6075) + (-4{\cdot}7477)\}$$

$$= \exp\{(1{\cdot}0274)(0{\cdot}273) - (4{\cdot}7477)\} = 1{\cdot}1E-2$$

procedures contained within this study to estimate an error of intention. The assumption is made that the identification and selection of the appropriate error of intention is achieved and that plant specific information is available and accurate. The site specific scenario and accompanying calculations are described below.

Personnel at plant X have a history of operating outside of procedures. In addition, general housekeeping, i.e. the match between piping and instrumentation diagrams, procedures and operator schematics and control room activities, is less than optimal. The safety culture, is also less than optimal. Many piping elbows are 'bagged' and tags are missing from important safety equipment. After performing a task analysis and systems walkdown, an HRA analyst recognizes the need for the HRA analysis to model the possibility of an error of intention occurring in which personnel make 'back of the envelope' calculations instead of using the formuli provided inside procedures.

In order to estimate the HEP for this error, the analyst first consults Table 2 and finds the specific HEP which best corresponds to the action sequence of interest. In the current example, error of intention no. 4 'Violate procedures—personnel devise their own formuli' closely corresponds to the scenario under consideration. This table presents extreme HEP estimates corresponding to best and worst case PSF scenarios. If only very general information was available about the PSFs at plant X, i.e. all that is known is that, in general, the PSFs are very favorable, then the HEP for use in the HRA would be determined by selecting the LB; or conversely, if the PSFs are very unfavorable, the UB.

Since we have relatively complete information available to us regarding plant X, the analyst can assess the PSFs there. The analyst rates the 11 PSFs at plant X on a Likert scale from 1 to 5, where 1 is least favorable and 5 is most favorable. Table 3 contains PSF ratings and can be used to better define the HEP estimate within the range provided by the UB and LB presented in Table 1.

The analyst then multiplies each of his or her own Likert ratings by each of the corresponding weights for these PSFs given in Table 1. Next these 11 sets of weights and ratings are summed. This gives the observed composite or weighted PSF score which represents the 'x' in the left hand side of eqn (1). The product of the ratings and the weights are presented in Table 4. For the present example, solving for this equation using UB = $4{\cdot}7E-2$, LB = $1{\cdot}6E-3$, and PSF composite = 257, the HEP for personnel at this site violating procedures and using their own formuli is equal to $1{\cdot}1E-2$, as shown in Table 4.

6 DISCUSSION AND SUMMARY

A method, INTENT, is reported whereby risk analysts may account for the influence of errors of intention in PRAs. Using a hypothetical example, it was possible to make use of a preliminary data set for 20 errors of intention that was tailored to represent the influence of 11 commonly referenced performance shaping factors. Lower and upper bound HEP data for each of the errors were generated by HRA and human factors analysts for best case and worst case performance shaping factors. A formuli was derived for mapping composite PSF scores onto the HEP scale. Preliminary review suggests that the method provides an interim mechanism to provide data which can serve to remedy a major deficiency present in all PRAs conducted to date; failure to account for rare, high consequence failures due to errors of intention committed during various plant evolutions. Since the upper and lower bound HEP data for errors of intention are derived from expert estimation they should be used judiciously until such time as they can be replaced with operations, i.e. field data. The reader is advised that the errors resulting from an initial data search using SNEAK method techniques have been categorized to reflect the data themselves and are not to be construed as replacing cognitive activities themselves such as goal setting, planning, analyzing, solving by analogy, etc. The categories (i.e. consequence, attitudes, response set, and resource dependency) do have a basis in the literature wherein they may be construed as influences upon performance.

Future research should attempt to (1) extend the variety of errors of intention identified in this paper, (2) determine whether the suggested PSF weights will hold true for these 'new' errors of intention, (3) apply the INTENT method in an HRA program, and (4) attempt to tie the data and categories to some high level cognitive theory. For example, the categories of response set and resource dependency may be reviewed for their compatibility with error shaping factors proposed for level II and III of the Generic Error Modeling System (GEMs) hierarchy developed by Reason.[19] Hopefully, a contribution of this present work will lie in enabling the HRA and PRA community to quantitatively account for an important aspect of the variability in systems performance. As a corollary to this research, effort needs to be taken to identify appropriate precursors to the various errors presented in this paper. The latter objective may be difficult to achieve as erroneous actions not leading to LERs are currently unnoticed, or if noticed, tend to go unreported. Secondly, if as the authors believe, the situations required to produce such errors are themselves infrequent, then longitudinal studies may be required in order to collect these data. In the interim, we are attempting to refine and validate the INTENT approach for quantifying decisionbased errors.

ACKNOWLEDGEMENTS

The authors wish to thank Dr S. Hill of the INEL Human Factors Research Unit and Dr Donald L. Schurman of Science Application International Corporation (SAIC) for review of the manuscript and to C. Gentillon for her insights regarding the sections on quantification. J. Auflick, O. Meyer and B. Hallbert are also offered much thanks for their participation and insights during the conduct of the SLIM sessions.

REFERENCES

1. Swain, A. D. & Guttman, H. E., *Handbook of Human Reliability Analysis with Emphasis on Nuclear Power Plant Applications.* NUREG/CR-1278, US Nuclear Regulatory Commission, Washington, DC, 1983.
2. Hollnagel, E., Ragnar, R. & Taylor, R. S., Human Reliability and the Reliability of Cognition, *Cognitive Reliability,* **4** December (1989) 1–12.
3. Johnson-Laird, P. N., In *Handbook of Cognitive Neuroscience,* ed. M. S. Gazzaniga, Plenum Press, New York, 1984, 249–74.
4. Reason, J., Managing the management risk: new approaches to organizational safety. Workshop on Managing New Technologies, Bad Homburg, Germany, 3–5 May 1990.
5. Kahneman, D., *Attention and Effort,* Prentice Hall, Englewood Cliffs, New Jersey, 1973.
6. Norman, D. & Bobrow, D., On data-limited and resource-limited processing. *Journal of Cognitive Psychology,* **7** (1976) 44–60.
7. Kahneman, D. & Tversky, A., Prospect Theory: An analysis of decisions under risk. *Econometrica,* **47** (1979) 263–91.
8. Woods, D. D., Roth, E. M. & Pople, H., Jr, Modeling human intention formation for human reliability assessment. *Reliability Engineering and System Safety,* **22** (1988) 169–200.
9. Yeh, Y. C. & Teng, Y. T., Quantification of dynamic human error probabilities in terms of sampled timing. *Nuclear Technology,* **84** (1989) 252–5.
10. Dougherty, E. M., Jr, Human reliability analysis—Where shouldst thou turn? *Reliability Engineering and System Safety,* **29** (1990) 283–99.
11. Ujita, H., A probabilistic analysis method to evaluate the effect of human factors on plant safety. *Nuclear Technology,* **76** (1987) 370–6.
12. Hahn, H. A., Blackman, H. S. & Gertman, D. I., Applying SNEAK analysis to the prediction of human error. *Reliability Engineering and System Safety,* **33** (2) (1991) 289–300.
13. Harbour, J. L. & Hill, S. G., *HSYS: A Methodology for Analyzing Human Performance in Operational Settings,* EG&G Technical Report, EGG-8806, EG&G Idaho Inc., Idaho Falls, Idaho, 1990.
14. Comer, K. & Donovan, M., NUREG/CR-4010, General Physics Corporation 1983. (In 1987, the INEL expanded the concept to include hardware component failure rate data and has proceeded with the implementation.)
15. Gertman, D. I., Gilmore, W. E., Galyean, W. J., Groh, M. R., Gentillon, C. D. & Gilbert, B. G., *Nuclear Computerized Library for Assessing Reactor Reliability Volume I: Summary Description and Volume V: The Data Manual.* (With W. J. Reece & G. B. Beers) NUREG/CR-4639, US NRC Washington, DC, February, 1989.
16. Comments regarding crew '*reluctance*' as a factor which should be accounted for in HRAs were voiced by John Wreathall August of 1989 during discussions with the senior author. He will, no doubt, be surprised to find himself so quoted out of context.
17. Whyte, G., Groupthink reconsidered. *Academy of Management Review,* **14** (1) (1989) 40–56.
18. Embrey, D. E., Humpherys, P., Rosa, E. A., Kirwan, B. & Rea, K., *SLIM MAUD: An Approach to Assessing Human Error Probabilities Using Structured Expert Judgement.* Vol. I. (Overview of SLIM-MAUD, NUREG/CR-3518 Vol. I.) US Nuclear Regulatory Commission, Washington DC, March, 1984.
19. Reason, J., In *New Technology and Human Error,* ed. J. Rasmussen, K. Duncan & J. LePlat. John Wiley and Sons, London, 1987.