The quantification of human fallibility

Paper presented to the BHSS Conference on "Human Factors in Managing Health and Safety"

Dr Linda J Bellamy
Four Elements Ltd, London

Abstract – Recent disasters have highlighted the problems of operational and maintenance errors, as well as the underlying management and organisational problems. This paper aims to illustrate some of the reasons for quantifying human error and the problems encountered in trying to do so. Rather than attempting to examine the subject in depth, the intention is to present an overview of some of the central issues.

What is human error?

Human beings, by virtue of their behavioural variability, have an enormous capacity to learn and to adapt to their environment. There are many work situations, however, where human activity in the work environment must be constrained because variability in behaviour cannot be tolerated by the system – for instance, in the operation of a hazardous chemical plant. These constraints may take the form of rules and procedures.

Human error might therefore be defined in the context of the person operating in an intolerant system. Human error is frequently given as a cause of failure in a system where human action (or inaction) has exceeded system tolerances and resulted in some undesirable consequence.

Some definitions of error examine internal human mechanisms. An example is Reason's 'slips' and 'mistakes' (REASON, 1987):

- *Slips* – departures of action from intention or execution failures (*eg.* attention slips in routine actions);
- *Mistakes* – errors in which the action may run according to plan, but where the plan is inadequate to achieve its desired outcome (*eg.* selecting an incorrect procedure due to an incorrect diagnostic inference.

Others provide a definition in terms of external modes of functioning, such as Swain and Guttmann's classification of incorrect human outputs (SWAIN and GUTTMANN, 1983):

- *Errors of omission*
  - omits entire task;
  - omits a step in a task.
- *Errors of commission*
  - selection error (selects wrong control, mispositions control or issues wrong command);
  - error of sequence;
  - time error (too early, too late);
  - qualitative error (too little, too much).

We should include in this list 'extraneous acts' or 'wilful violations'. These are actions taken by an operator outside the usual procedures, as happened at Chernobyl.

RASMUSSEN (1987) points out that these external modes of malfunction are the first elements of man-system mismatch encountered when backtracking a course of events to identify the causes of an unacceptable incident (*eg.* the release of a hazardous chemical).

It is these external modes of behaviour that are generally the focus of quantification. However, the classification of error in these terms is severely limiting without consideration of causes. It is important to understand why human errors occur, not only in order to quantify them but also to reduce their likelihood of occurrence if unacceptably high.

Why quantify human error?

It was pointed out in a recent letter to *Loss Prevention Bulletin* (BOND, 1990) that this could be the 'Human Factors' decade. Human error in industries where human and hardware failures can cause loss of life is a very topical subject. Many recent disasters have highlighted the problems of operational and maintenance errors, and also the underlying management and organisational problems (*eg.* King's Cross, Zeebrugge, the Challenger space shuttle, Clapham Junction and Chernobyl). It is important to estimate how likely accidents with such serious consequences are and, particularly, how they could occur.

Human error as a direct cause of failure is a significant contributor to accidents. FIGURE 1 (from BELLAMY et al, 1989) shows the percentage contribution of different direct causes of pipework failures, expressed as a percentage of known causes. Operator error (for operations, maintenance, testing, *etc.* activities) accounted for 31% of known causes and was the largest

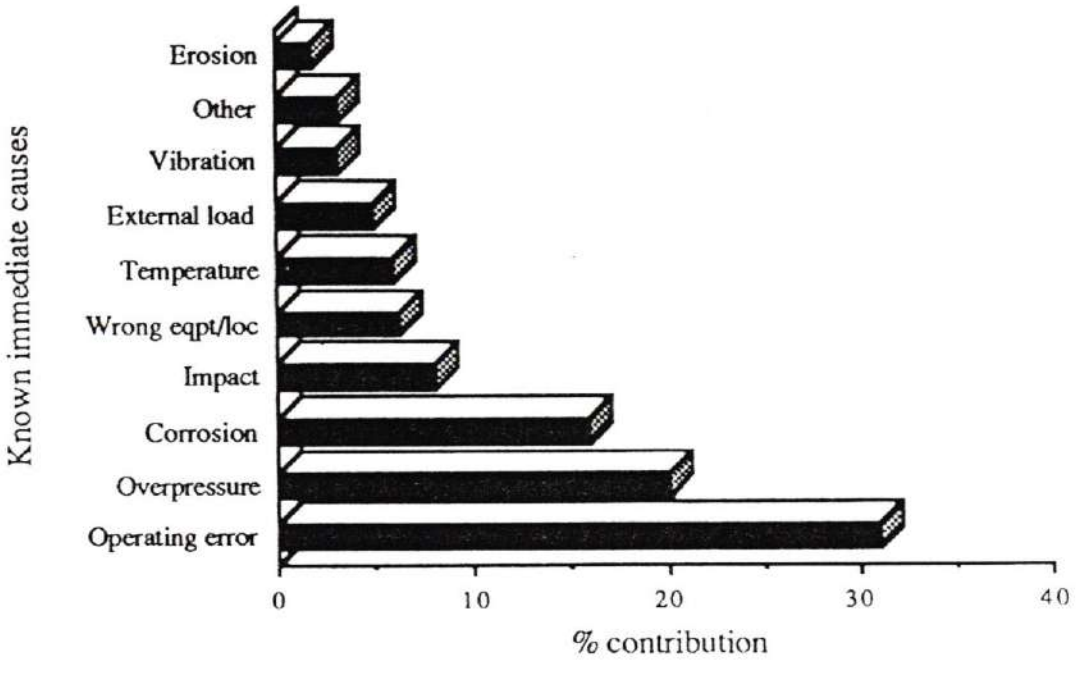

Figure 1

Breakdown of known direct cause contributions to pipework failure

contributor. These errors were principally failures to clear or isolate pipework from hazardous materials before working on it (24% of operator error causes) and incorrect setting of equipment status, such as valves and pumps (17% of operator error causes).

Some systems have several lines of defence against the realisation of hazards such that a single error should not produce disastrous consequences. However, different combinations of human errors and hardware failures can form adverse configurations such that there are various routes to an accident and its consequences. Sometimes errors can become latent in a system because they have no immediately observable adverse consequences – thereby contributing to accidents 'waiting to happen'.

To give an example of such latency, a study of 17 computerised process control system incidents (BELLAMY and GEYER, 1988) indicated that, in 10 of the incidents, operators had been working on inadequate or incorrectly supplied information. For only two of the incidents had operators actually made errors in diagnosing abnormal conditions when the supply of information was adequate. Latent information problems included:

- man-machine interface not displaying actual plant status;
- errors in installing instrumentation;
- alarm disabled during maintenance;
- failure to supply an alarm in the design.

Because the human operator can act as a line of defence against failure consequences – for example, by detecting and correcting a failure which could lead to a release of a hazardous chemical – it is as important to design, review and test the human support aspects of the system at least as well as is done for hardware and other engineered aspects. Human errors when a system is in an abnormal state can be disastrous – for instance, the Three Mile Island nuclear accident in 1979, when operators cut off the emergency coolant, and the Boeing 737 crash on the M1 motorway near Kegworth in January 1989 when the wrong engine was shut down.

In summary, then, one reason for quantifying error is that it helps to identify where, for hazardous operations, the potential for human errors with adverse consequences must be addressed in improving safety and reliability. Another reason is that, as a contributor to failure or inability to control an incident, human error probability affects the likelihood of outcomes of an incident and this information is an important contributor in the overall assessment of whether a system is acceptable.

Human error and risk assessment

In attempting to quantify the human errors that may occur in a potentially hazardous system, for the purposes of risk assessment one is trying to provide data on this human contribution to the risk picture, where the risk is represented by the likelihood of the realisation of the hazards together with the possible outcomes of different degrees of severity.

If the likelihood of a train driver passing a stop signal is assessed as being high, one might look for possible ways of improving the system to reduce this likelihood, such as:

- improved visibility of the signal (*eg.* location, brightness, size);
- additional warning information (*eg.* audible signals);
- driver selection and training.

These are all 'human factors' approaches. However, if it is not possible to reduce the error likelihood significantly, the designer might examine ways of reducing the consequence likelihood, such as the introduction of automatic devices for stopping or re-routeing the train after it has passed the stop signal.

Usually the design of technology and procedures in hazardous industries address the reduction of the likelihood of both causes and consequences of failure. Risk must be reduced to a level which is 'acceptable'. If risk outweighs benefit or the cost of reducing risk is too high then a particular activity might be rejected as being unacceptable.

Risk is the likelihood that an event will happen and lead to adverse consequences – for example, the likelihood per car journey of having a tyre blowout and dying as a result. However, the car driver who voluntarily takes a risk does not make a formal assessment of how likely she/he is to die during a particular journey although, if asked, may be able to express her/his evaluation quantitatively (*eg.* a 1 in 10,000 chance of dying). Perception of risk may be based on a number of factors such as ...

- annual road death figures;
- judgement of driving ability;
- knowledge of route;
- evaluation of car condition;
- hazards encountered on route taken;
- *etc.*

... rather than a detailed analysis and synthesis of the many components of the risk picture, some of which are shown in FIGURE 2.

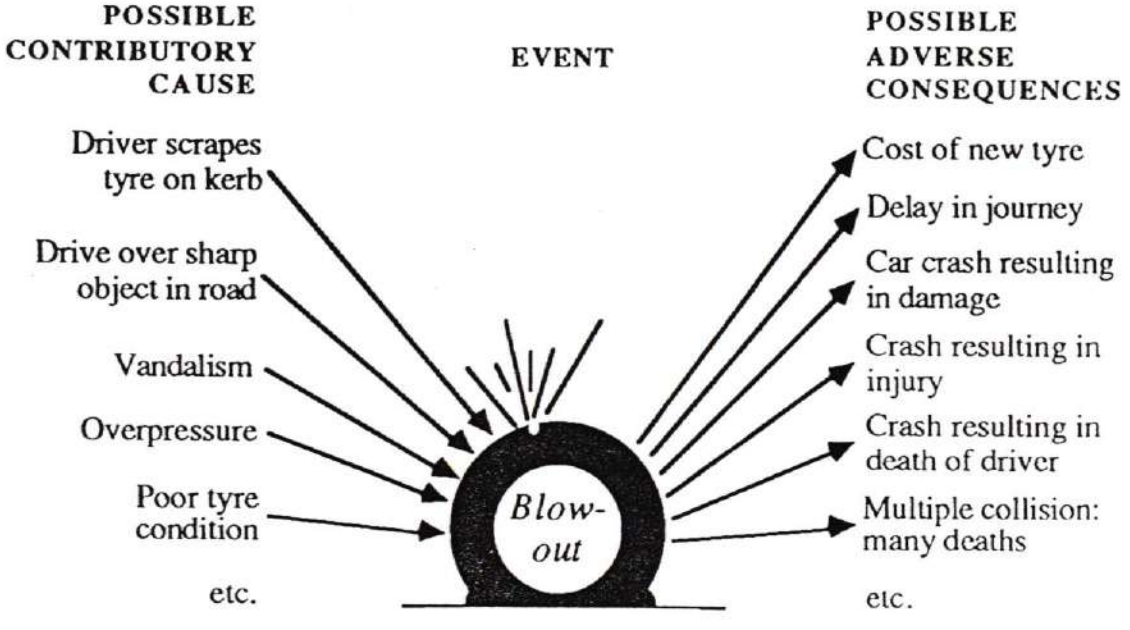

Figure 2

Part of the risk picture for a tyre blowout

A formal assessment would examine possible events and their consequences in a systematic way, ascribing numerical values to event and outcome likelihoods by scrutible procedures.

An example of a formal assessment is illustrated in FIGURE 3 (from BELLAMY et al, 1986), which shows a fault tree for the event "failure to launch lifeboat at first attempt" from an offshore platform. This is a logic diagram showing the interrelationships between all of the contributory causes (not identified in diagram). The human errors are highlighted. This form of analysis attempts to look at the total system. It treats the human operator as a component of the system, demands all the failure modes of the operator to be considered and ultimately requires the estimation of the probability of these modes.

Risk assessment is a scrutible method which is generally applied to 'involuntary' risks (*eg.* the risk posed by a chemical plant), and which quantifies risk and makes comparison with a criterion. Decisions can then be made as to whether the risk is 'acceptable' or whether the selection of a particular design or procedure is preferable to another. A formal risk assessment would produce quantified answers to questions concerning, for example, the risk to life or the environment from the siting of a particular chemical installation.

In risk assessment, the expression of likelihood as a number (probability or frequency) is useful because it allows comparisons to be made easily, for example, with a performance standard. However, the numbers are not exact. Uncertainty exists to a degree that is dependent upon knowledge of:

- past events;
- causal relationships.

This leads to the question of whether human error can be quantified.

Is it possible to quantify human error?

In answering this question, the first response must be that it is possible to produce numbers to indicate the likelihood of error in tasks with known or assumed characteristics. A great deal of effort has gone into the development of Human Reliability Assessment (HRA) techniques over the past decade. The results of HRA are expressed in the form of human error probabilities or rates:

$$\text{Human error probability (HEP)} = \frac{\text{Number of errors}}{\text{Number of opportunities for error}}$$

$$\text{Human error rate} = \frac{\text{Number of errors}}{\text{Total task duration}}$$

HEPs are typically in the range of 1 to 0.00001. Note, however, that estimates of task success must take account of error recovery. For example, HEPs for using a calculator are high, but there are opportunities for recognising and correcting errors.

However, formal assessments requiring data on human error probabilities suffer from a lack of empirical data such as experimental, simulator or historical data. Therefore it is often necessary to generate data for new tasks and technologies using specially developed assessment methods.

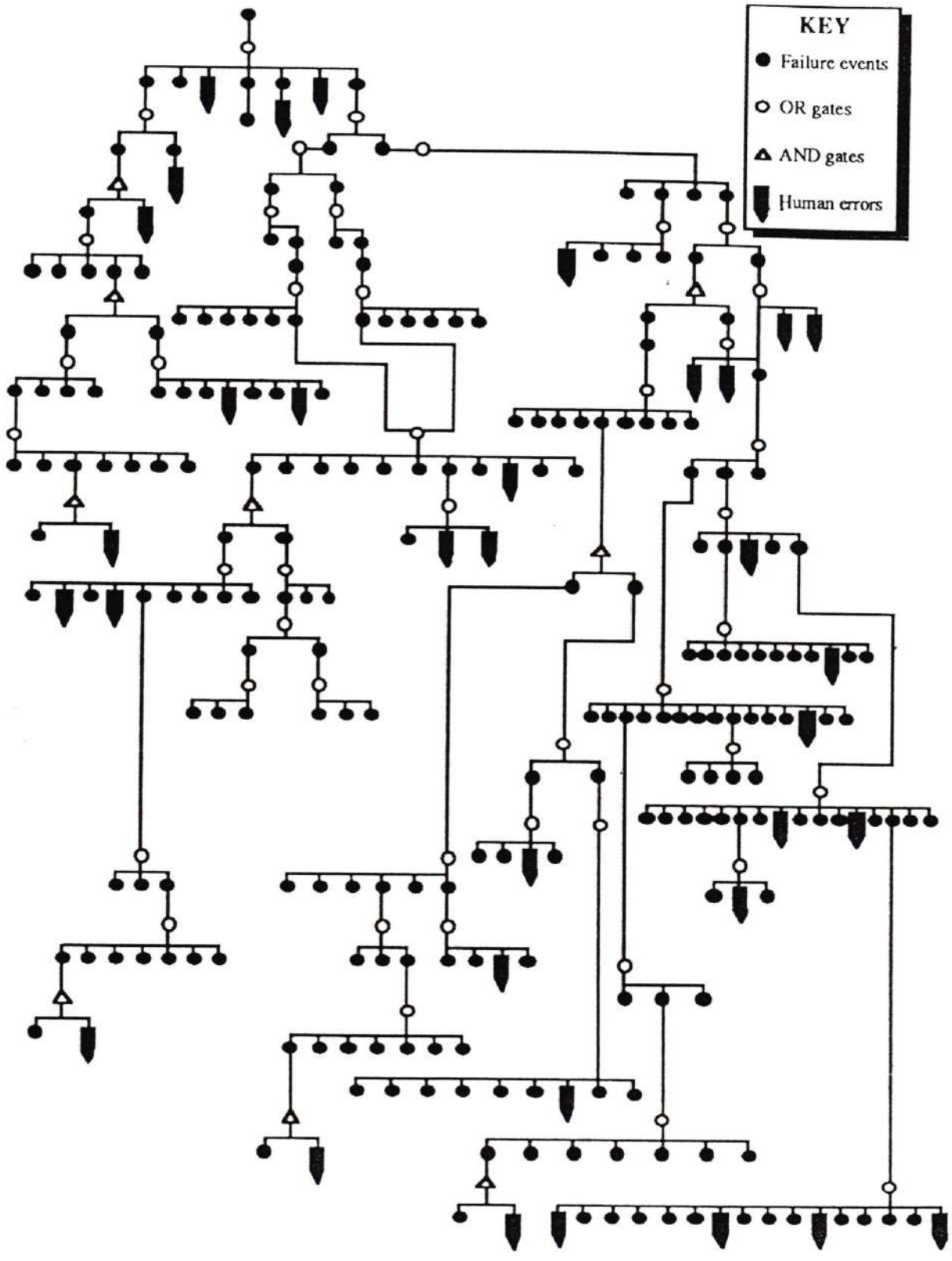

Figure 3
Schematic of a fault tree for lifeboat launching showing human error contributions

HRA techniques have some or all of the following characteristics:

- Identification of relevant tasks performed (or to be performed) by operators;
- Representation of each task by some method or model (*eg.* decomposition of the task into its principal components by task analysis methods to determine opportunities for error and error recovery);
- Identification of conditions which affect error probability or rate (performance shaping factors);
- Use of data derived from historical records or judgements (and usually both);
- The prescription of error reduction strategies..

It is not the intention here to go into detail about the different kinds of techniques. This task has recently been accomplished by the Human Reliability Assessment Group (HUMPHREYS, 1988). However, it is useful to consider the different kinds of approach. These tend to fall into one of three categories:

- **Data bank approaches** – which provide sets of error probability data and performance shaping factor multipliers along with the modelling technique – for example:
  - *Technique for Human Error Rate Prediction (THERP)* developed by SWAIN and GUTTMANN (1983) for the nuclear industry.
  - *Human Error Assessment and Reduction Technique (HEART)*, which explicitly considers performance shaping factors such as experience, overload and information problems (WILLIAMS, 1988).
- **Time dependent models** – for example:
  - *Human Cognitive Reliability (HCR)* technique for assessing operating team reliability under time constrained emergency conditions (HANNAMAN et al, 1985). Here, data are provided on the probability of failing to diagnose and respond to an abnormal event within time T after a signal indicating abnormality.
- **Expert judgement approaches** – for example:
  - *Paired comparisons (PC)* (HUNNS and DANIELS, 1980). Here, experts compare many pairs of tasks, at least two of which have known HEPs. The latter are used to calibrate a scale of tasks in terms of relative error likelihood.
  - *Absolute probability judgement (APJ)* or *Direct Estimation* (COMER et al, 1984. Methods range from simple guessing to the use of a group of experts.
  - *Influence Diagram Approach (IDA)* (PHILLIPS et al, 1985), which models and weights the combined influence of performance shaping factors as well as using direct estimation of HEPs.

However, the quantification of human error poses problems for a number of reasons:

- The modelling of error causes has only recently been influenced by cognitive psychology. Although this has provided an insight into error causal mechanisms and error classification, there are no causal models which can be used to generate HEP data without the use of expert judgement or historical data.

- Incident records rarely describe the performance shaping factors (PSFs) or conditions under which the human errors were made. This makes it difficult to generalise from failure data for specific tasks to others that are similar. PSFs are important because they have a significant effect on human error probability. FIGURE 4 shows how extensive the range of PSFs for consideration might be (although in HRA one usually considers only the major ones).

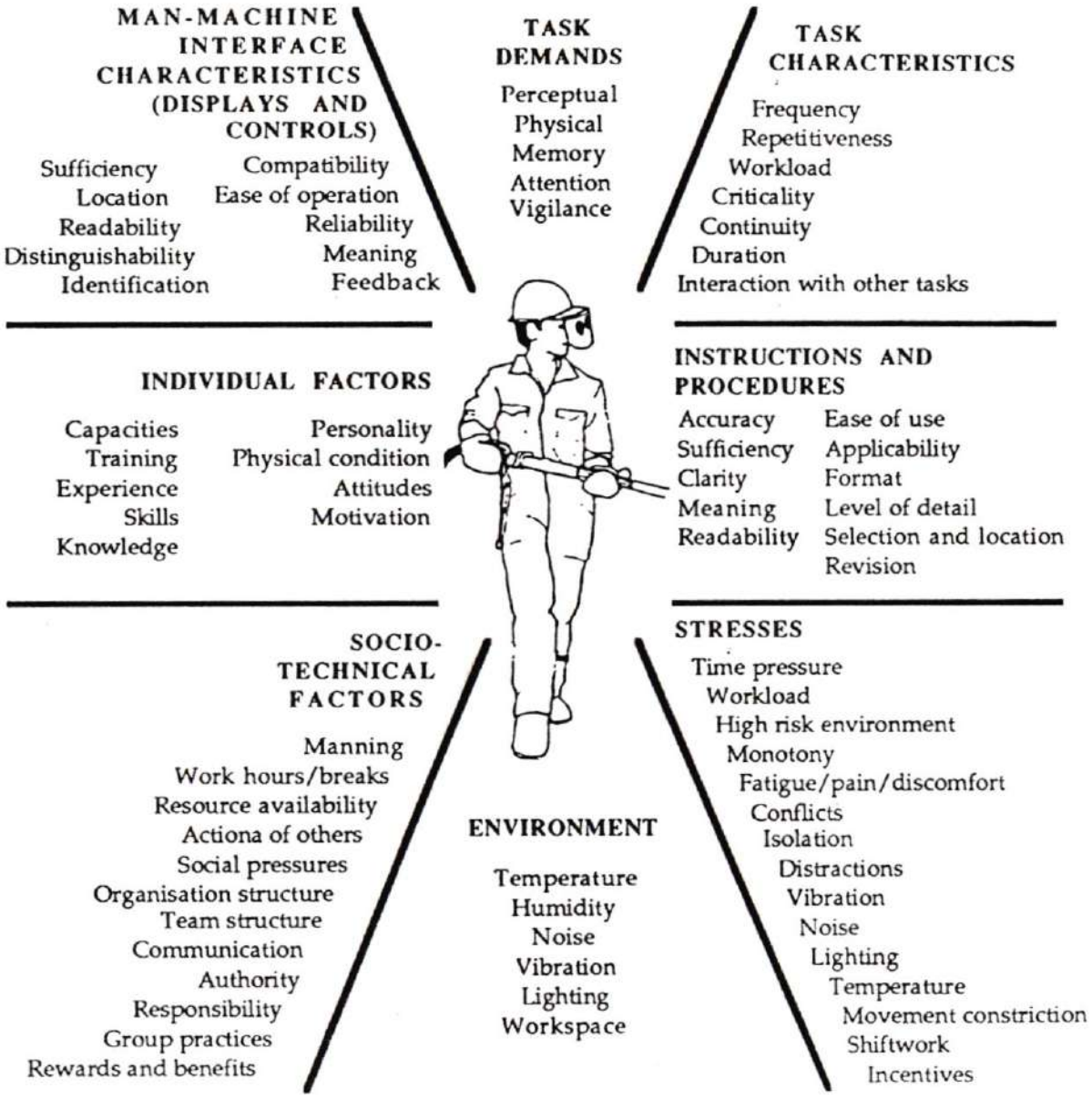

Figure 4
Performance shaping factors

- Incident records only reflect errors which have been identified and which resulted in some notifiable consequence. They do not record either opportunities for error or error frequencies with no consequence (*e.g.* because of error recovery) unless there is a good 'near miss' reporting scheme. It is not possible to determine, therefore, what the true error rate is although attempts can be made by estimating opportunities for error.

- Techniques which provide data, such as THERP, do not provide the original data on which the numbers are based.
- There may be bias in making expert judgements and in judging error probabilities for new designs of systems which have not yet been operated, although progress has been made in 'structuring' the judgements in order to reduce the bias.
- Many of the quantification techniques have not been validated and can suffer from variability in analysts' modelling and judgements.

Conclusion

Bearing in mind the limitations described above, some attempt at quantification must be made in order to assess the human contribution to risk. It is also necessary in order to be able to prioritise design improvements which will reduce this risk. Work on the quantification of error must continue in order to improve safety and reliability. Interest from industry and regulators is already extending to the even more complex problem of how to quantify the effects of management quality on risk (BELLAMY et al, 1990).

Quantification is useful because:

- the process of quantifying requires a systematic examination of a system;
- it provides data necessary for decision-making such as prioritising design improvements;
- it prompts the identification and understanding of differences which might otherwise be overlooked;
- it is part of the process of risk reduction.

References

BELLAMY LJ and GEYER TAW (1988) Addressing human factors issues in the safe design and operation of computer controlled process systems. **In** SAYERS BA (ed) *Human factors and decision making*. Elsevier: London

BELLAMY LJ, GEYER TAW and ASTLEY JAA (1989) *Evaluation of the human contribution to pipework and in-line equipment failure frequencies. (HSE Contract Research Report Nº 15/1989)* Health & Safety Executive: Bootle

BELLAMY LJ, GEYER TAW, WRIGHT MS and HURST NW (1990) *The development in the UK of techniques to take account of management, organisational and human factors in the modification of risk estimates.* Paper presented at the American Institute of Chemical Engineers' Spring National Meeting "Emerging Technologies", 18-22 March 1990, Orlando, Florida

BELLAMY LJ, KIRWAN B and COX RA (1986) *Incorporating human reliability into probabilistic risk assessment.* Paper presented at the 5th International Symposium on "Loss Prevention and Safety Promotion in the Process Industries, Societé de Chimie Industrielle, Paris

BOND J (1990) What next? *Loss Prevention Bulletin* **92**: 33-34

COMER MK, SEAVER DA, STILLWELL WG and GAI DY CD (1984) *Generating human reliability estimates using expert judgement, vol. 1 Main Report (NUREG CR-3688/10F2.5 and 84-7115, GP-R-213022)*

HANNAMAN GW, SPURGIN AJ and LUKIC Y (1985) *A model for assessing human cognitive reliability in PRA studies.* IEEE 3rd Conference on Human Factors and Nuclear Power Plants, Monterey, Calif., 23-27 June 1985

HUMPHREYS P (ed) (1988) *Human reliability assessors guide (RTS 88/95Q).* Safety and Reliability Directorate

HUNNS D and DANIELS BK (1980) The method of paired comparisons and the results of the paired comparisons consensus exercise. **In** *Proceedings of the 6th Advances in Reliability Technology Symposium, vol. 1 (NCSR R23)* National Centre of Systems Reliability: Culcheth, Warrington

PHILLIPS LD, HUMPHREYS P and EMBREY DE (1985) Appendix D: A socio-technical approach to assessing human reliability (STAHR). **In** SELBY D *Pressurized thermal shock evaluation of the Calvert Cliff Unit 1 Nuclear Power Plant (Research report on DOE Contract 105840R21400)* Oak Ridge National Laboratory: Oak Ridge, TN

PHILLIPS LD and EMBREY DE (1985) Appendix E: Quantification of operator actions by STAHR. **In** SELBY D *Pressurized thermal shock evaluation of the Calvert Cliff Unit 1 Nuclear Power Plant (Research report on DOE Contract 105840R21400)* Oak Ridge National Laboratory: Oak Ridge, TN

RASMUSSEN J (1987) The definition of human error and a taxonomy for technical system design. **In** RASMUSSEN J, DUNCAN K and LEPLAT J (eds) *New technology and human error.* John Wiley & Sons: Chichester

REASON J (1987) Generic error-modelling system (GEMS): a cognitive framework for locating common human error forms. **In** RASMUSSEN J, DUNCAN K and LEPLAT J (eds) *New technology and human error.* John Wiley & Sons: Chichester

SWAIN AD and GUTTMANN HE (1983) *Handbook of human reliability analysis with emphasis on nuclear power plant applications (NUREG/CR-1278).* Sandia National Laboratories, Alburquerque, New Mex. Report prepared for US Nuclear Regulatory Commission: Washington DC

WILLIAMS JC (1988) A data-based method for assessing and reducing human error to improve operational experience. **In** *Proceedings of IEEE 4th Conference on Human Factors in Power Plants, Monterey, Calif. 6-9 June 1988*