PII:S0003-6870(98)00010-6

Applied Ergonomics Vol. 29, No. 3, pp. 157-177, 1998

0003-6870/98 $19.00 + 0.00

Human error identification techniques for risk assessment of high risk systems—Part 1: review and evaluation of techniques

Barry Kirwan*

Industrial Ergonomics Group, School of Manufacturing and Mechanical Engineering, University of Birmingham, B15 2TT, UK

(Received 29 March 1996; accepted 8 January 1998)

This is the first in a two-part series of papers dealing with the area of assessing human errors in high risk complex systems. This first paper outlines thirty-eight approaches of error identification, categorising them into types of error identification approach. The paper then reviews these techniques with respect to a broad range of criteria. Viable and non-viable techniques are identified. Trends and research needs are also noted. The second paper proposes a framework or tool-kit approach to Human Error Identification, and presents a prototype methodology to show what such a framework approach would look like in practice, for the nuclear power domain. © 1998 Elsevier Science Ltd. All rights reserved.

Keywords: human error, risk assessment, techniques, prediction

Introduction

This review builds upon and extends an earlier review of twelve human error identification techniques (Kirwan, 1992a, 1992b). A good deal of introductory information on human error analysis is given in these previous two papers, and is only briefly summarised here. The reader unfamiliar with this territory is therefore referred to these articles or to Kirwan (1990, 1994), or Embrey *et al.* (1994) which also review a number of techniques. Essentially, human error identification (HEI) is usually a part of Human Reliability Assessment (HRA), which determines the impact of human error and error recovery on a system. Such predictions feed into risk assessments known as Probabilistic Safety Assessments (PSAs: Green, 1983; Cox and Tait, 1991) which determine the risk of systems such as nuclear power plants, chemical plants, etc. from all potential risk causes, including human error. This risk assessment process of determining whether a plant is safe to operate or build, or whether it should be altered, shut down, or cancelled, is critically dependent on human error identification.

The process of error identification begins with deciding the scope of the analysis. A principal decision is which operator involvements to consider: whether to consider only emergency events, or whether to consider misdiagnoses, maintenance errors, or even rule violation errors. Once the scope of the assessment and the critical tasks to be assessed have been identified, the next phase is *task analysis*, determining how the operations should proceed, a necessary step prior to considering how they may fail. Once a task analysis representation has been achieved, the analyst or assessor may then begin to consider what could go wrong. The assessor uses techniques of error identification to interrogate the formal description (e.g. a task analysis) of what the operator should do, and thus identify what could go wrong. There are three major components to an error:

- *External Error Mode* (EEM) the external manifestation of the error (e.g. closed wrong valve)
- *Performance Shaping Factors* (PSF) which influence the likelihood of the error occurring (e.g. quality of the operator interface, time pressure, training, etc.)
- *Psychological Error Mechanism* (PEM) the 'internal' manifestation of error (how the operator failed, in psychologically meaningful terms, e.g. memory failure, pattern recognition failure, etc.)

The external error mode is the minimum that needs to be identified, since this is what will 'appear' in the risk assessment (e.g. in the PSA fault and/or event trees). Performance Shaping Factors and Psychological Error Mechanisms are desirable, however, as they aid in the quantification process and are useful when determining

*Barry Kirwan is currently Head of Human Factors, ATMDC, National Air Traffic Services, Bournemouth Airport, BH23 6DF, England. This papers was prepared while on sabbatical in the Safety Science Group, Technical University of Delft, Netherlands.

how to reduce the likelihood of an error via error reduction mechanisms.

Once errors have been identified, error recovery potential can be considered, as well as the consequences of the identified error. These qualitative considerations of the error identified amount to *human error assessment* (identification, error recovery consideration, and consequence determination). If the likelihood or probability of the error is then calculated this amounts to a full *Human Reliability Assessment*. Furthermore, ways of reducing the likelihood of the error, or its impact on the system, can be specified, and this is known as Error Reduction Analysis (ERA). These are the major steps in the HRA process. Before continuing with the types of errors that are identified, and the review of the techniques themselves, it is worth briefly examining the relationship between error identification and Ergonomics.

Ergonomics and error identification

Error reduction analysis is clearly (along with task analysis) the most relevant aspect of HRA to Ergonomics. Error reduction recommendations are most often readily apparent at the human error assessment phase, i.e. before quantification, and so are not dependent on quantification. This point is made since some ergonomists ignore the potential benefits of error identification in HRA because of their reluctance to accept HRA's quantification stance, but in fact quantification is not a necessary function of HRA, unless being used expressly for PSA purposes. This means that ergonomists can utilise human error assessment approaches and gain valuable error reduction insights, without having to quantify human error likelihoods or probabilities. Therefore, although the techniques reviewed in this paper may have been developed primarily for supporting HRA and PSA, some of them may be of more general interest to ergonomists who have no interest in HRA, but who wish to predict potential error modes with systems or products. Having made this statement, the rest of this paper (and its sequel) will nevertheless focus on HRA-related applications, since that is the main focus of this technique review and evaluation.

Error types of interest in risk assessment

The types of error contribution of interest in risk assessments for the high risk industries have been determined in earlier reviews (Reason, 1990; Rasmussen *et al.*, 1981; Kirwan, 1993, 1994), and are themselves evident from the Human Error Identification (HEI) techniques that have been developed, sometimes with the specific purpose of identifying a certain error type. The major error types (Kirwan, 1993) of interest are as follows:

- *Slips and lapses* (also called action execution errors) —These errors are the most predictable, and are usually characterised by being simple errors of quality of performance (e.g. too much, or too little force applied), or by being omissions, or sequence errors (task steps carried out in wrong sequence).
- *Cognitive errors: diagnostic and decision-making errors* —These relate to a misunderstanding, by the operators, of what is happening in the system (usually due to insufficient operator support [design; procedures; training], given the specific scenario occurrence). Such errors are of continuing concern due to their ability to alter accident progression sequences and to cause dependencies (in a failure sense) between redundant and even diverse safety and back-up technical systems. This type of error includes misdiagnosis, partial diagnosis and diagnostic failure.
- *Maintenance errors and latent failures*—Most maintenance failures will also be due to slips and lapses, but in maintenance and testing activities, which may lead to immediate failures or to latent failures (failures whose impact is delayed, and whose occurrence may be difficult to detect prior to an accident sequence). Most PSAs make assumptions that maintenance failures are implicitly included in component and system availability data, which necessarily include maintenance error contributions. However, it is less clear that such maintenance data used in the PSA can incorporate the full impact of latent failures (errors which lay dormant until a system is used).
- *Errors of commission*—An Error of Commission (EOC), is one in which the operator does something that is incorrect and also unrequired. Examples are where a valve which should be locked open is found to be locked closed. Such errors can arise due to carrying out actions on the wrong components, or can be due to a misconception, or to a risk recognition failure. These EOCs are becoming of increasing concern recently, for three reasons: firstly they do appear to happen, even if rarely; secondly, they can have a large impact on system risk; and thirdly, they are very difficult to identify (and hence anticipate and defend against) in the first place. This means that they may therefore be underestimated in terms of their contribution to risk in the PSA, and may not even be represented in the PSA.
- *Rule violations*—There are two main types (Reason, 1990): the '*Routine*' rule violation where the violation is seen as being of negligible risk and therefore the violation is seen as 'acceptable' and even a necessary pragmatic part of the job; and the 'Extreme' violation, where the risk is largely understood as being real, as is the fact that it is a serious violation. The latter types of violation are believed to be rare. PSAs rarely include violations quantitatively, but the current trend is at least acknowledging their credibility. The concern with rule violations is similar to that for cognitive errors - rule violations are relatively unexpected and can lead to failure of multiple safety systems and barriers.
- *Idiosyncratic errors*—These are concerned with errors due to social variables and the individual's current emotional state when performing a task. They are the result of a combination of fairly personal factors in a relatively unprotected and vulnerable organisational system. HRA practitioners are often asked the question, 'how can you predict what an operator might do, where that operator has had a really bad row with his/her spouse that morning?', or 'what if two operators won't talk to each other because of a dispute?'. Essentially some accidents do fall into this category (e.g. possibly the Aberfan disaster; see Kirwan 1994), and they are extremely difficult to predict, as they relate to covert social factors not obvious from a formal examination of the work context. These errors are of particular concern where, for example, a single individual has the potential to kill a large number of persons (e.g. transportation systems). They are not dealt with in PSA or HRA.
- *Software programming errors*—These errors are of importance due to the prevalence of software-based control

systems required to economically control large complex systems (e.g. using a Distributed Control System [DCS] or a Supervisory Control and Data Analysis [SCADA] system). They are also important in other areas (e.g. navigational software in aviation) and for any safety critical software applications generally. Typically there are few if any techniques applied which predict human errors in software programming. Instead, effort is spent on verifying and validating software to show it is error-free. Unfortunately complete and comprehensive verification of very large pieces of software is intractable due to software complexity and interactiveness.

Both the areas of idiosyncratic errors and software programming errors are very ill-served by HEI techniques at present, and such errors are not predicted at all in PSAs. These two sources of error are therefore areas for future research.

Human error identification approaches

Twelve human error identification techniques were reviewed in Kirwan (1992a, b). Since that time, a number of new approaches have been published, and at the time of this review a total of thirty-eight techniques were identified. Some of these are prototypical in nature, and not ready to be used in real risk or ergonomics assessments. Nevertheless, since the objective of this work was partly to ascertain the range of functions of human error identification approaches, such relatively immature techniques have been included in the review process.

Many of the approaches (about half) have only 'appeared' in the last five years, and these are largely prototypes, still in their development phase. Some of these will probably not receive further funding and so will disappear, failing to mature into applicable techniques. However, these have been included because of the current international interest in HEI, and the realisation of its importance. This current number of HEI techniques on the market therefore reflects the concern for the improvement of HEI in HRA/PSA. Therefore, even if certain techniques do 'disappear' into history earlier than others, these contemporary approaches should be reviewed at this stage to see their intended directions, as these may yield useful insights and ideas for the hazardous industries. Good potential ideas are sometimes discontinued for reasons other than scientific ones.

The low maturity of many of the techniques also means they are relatively under-published, often being published in short conference articles rather than scientific journals, making it frequently difficult to fully determine the detailed rationales, functioning, and full capabilities of the techniques. There may therefore be some inaccuracies or misinterpretations by the author as to details of the techniques, as a function of sparse information.

The techniques identified and reviewed are shown in *Table 1*.

Table 2 attempts to show the basic relationships between the techniques as they have developed over time. Five broad classifications have been used to show the techniques' general orientation or form:

1. *Taxonomies*—Many techniques (e.g. THERP) tend to be taxonomic to some degree (i.e. they contain error taxonomies), but these techniques are largely solely taxonomic in nature. They offer checklists of error modes, and the reliance is placed on the analyst to interpret them in the context of interest.
2. *Psychologically based tools* — These are tools that rely on an understanding of the factors affecting performance. This group is particularly characterised by tools that consider error causes (PSF) and/or error mechanisms (PEMs). The classic technique here is the SRK framework, parent to a number of techniques which have borrowed from its conceptual framework to some degree. Some techniques in this category are clear derivatives of SRK, like the SHERPA family of tools, and others remain psychological in approach but are either not reliant on SRK (e.g. GEMS; TAFEI; PHECA), or are semi-independent from it (e.g. HRMS).
3. *Cognitive modelling tools*—These are tools that try to model cognitive aspects of performance, either in terms of relationships between knowledge items relating to symptoms of events (for diagnostic reliability assessment) (IMAS), or in terms of how various factors will affect cognitive performance aspects of the task (CREAM). This domain is perhaps the least mature of the human error analysis approaches, but also perhaps the most interesting, as it is an attempt to combine cognitive psychology, the currently dominant paradigm in psychology, with a human reliability attitude.
4. *Cognitive simulations*—These are generally computer simulations of operator performance (e.g. CES; COSIMO). This is the most sophisticated human error identification area, often relying on expert system-type frameworks to predict performance and error. Most of these tools are aimed at modelling and predicting cognitive performance rather than psycho-motor, as it is felt that the former is both the more important and more dominant contributor to risk in complex systems.
5. *Reliability-oriented tools*—These stem from the reliability approaches that have proven their worth with non-human reliability problems: principally Hazard and Operability Study (HAZOP), Failure Mode and Effect Analysis (FMEA), and Event Tree Analysis (see Green, 1983). The HEA tools in this domain are therefore either grafted onto their reliability parent framework (e.g. Human HAZOP; HEMECA; or they are adaptations of the original concept (e.g. PREDICT; COGENT); or else they are clearly focused by PSA concerns onto specific human error issues required for PSA integrity and insight (EOCA; PRMA; ATHEANA). The above groupings and *Table 2* provide a broad 'family tree' structure or framework for the techniques. This structure is inevitably subjective and based on the author's biases and preconceptions, but is hopefully useful to the reader to group what is otherwise an large and somewhat confusing set of disparate techniques and approaches. *Table 2* also notes that some techniques could appear in more than one column, but the intent of the table is to show that certain relationships, some weak and some strong, clearly exist, and have affected the scientific evolution of these methods. The table also shows where new approaches or ideas have appeared, which add to the breadth of human error identification as a whole, e.g. affordance approaches such as TAFEI; tools aimed at

error of commission analysis, from SNEAK to ATHEANA; and violation assessment tools such as PREDICT. The table also shows how certain approaches have remained popular, such as the SHERPA-style of approach, and more generally the taxonomic approach.

Finally, it can be used to show where current developmental interest lies, in areas such as error of commission and violation analysis, cognitive simulations (though interest and funding may be waning in this area), and more cognitive psychology-based techniques such as CREAM. These latter areas together constitute an increased interest in areas of human performance that are as difficult to predict as they are important in risk terms. Perhaps more importantly, they represent a shift towards focusing on the *context* in which the operator works, as the primary driver of human performance and error. There has been a real shift towards studying context in HRA, rather than simply basing error predictions on rather sterile descriptions of operational procedures back in the luxury of a consultancy office, rather than at the plant of interest. This shift will enhance the credibility of HRA, and its effectiveness and relevance to real installations being assessed, and will lead to more tangible and relevant and useful error reduction mechanisms for real systems. It is therefore a welcome shift, though it means that error prediction becomes harder due to the level of complexity of performance that is being analysed.

Table 1 Human error analysis techniques reviewed

HAZOP*	HAZard and Operability Study technique (Kletz, 1974)
SRK*	Skill, Rule and Knowledge-based behaviour model (Rasmussen *et al.* 1981)
CMA*	Confusion Matrix Analysis (Potash *et al.* 1981)
Murphy Diagrams*	(Pew *et al.* 1981)
THERP*	Technique for Human Error Rate Prediction (Swain and Guttmann, 1983)
DYLAM	DYnamic Logical Analysing Methodology (Amendola *et al.* 1985)
SHERPA*	Systematic Human Error Reduction and Prediction Approach (Embrey, 1986a)
IMAS*	Influence Modelling and Assessment System (Embrey, 1986b)
GEMS*	Generic Error Modelling System (Reason, 1987b; 1990)
PHECA*	Potential Human Error Causes Analysis (Whalley, 1988)
CADA*	Critical Action and Decision Approach (Gall, 1988)
TALENT	Task Analysis-Linked EvaluatioN Technique (Ryan, 1988)
HEMECA	Human Error Mode, Effect and Criticality Analysis (Whittingham and Reed, 1989)
HRMS*	Human Reliability Management System (Kirwan, 1990)
CES*	Cognitive Environment Simulation (Woods *et al.* 1990)
INTENT	[not an acronym] (Gertman, 1991)
SNEAK	[not an acronym] (Hahn and de Vries, 1991)
COMET	COMmission Event Trees (Blackman, 1991)
INTEROPS	INTEgrated Reactor OPerator System (Schryver, 1991)
TOPPE	Team Operations Performance and Procedure Evaluation (Beith *et al.* 1991)
TAFEI	Task Analysis For Error Identification (Baber and Stanton, 1991)
COSIMO	COgnitive SImulation MOdel (Cacciabue *et al.* 1992)
PREDICT	PRocedure to Review and Evaluate Dependency In Complex Technologies (Williams and Munley, 1992)
SCHEMA	Systematic Critical Human Error Management Approach (Livingston, *et al.* 1992)
PHEA	Predictive Human Error Analysis technique (Embrey, 1993)
TEACHER/SIERRA	Technique for Evaluating and Assessing the Contribution of Human Error to Risk [which uses the] Systems Induced Error Approach (*Embrey*, 1993)
COGENT	COGnitive EveNt Tree (Gertman, 1993)
CREWSIM	CREW SIMulation (Dang *et al.* 1993)
ADSA†	Accident Dynamic Sequence Analysis (Hsueh *et al.* 1994)
PRMA†	Procedure Response Matrix Approach (Parry, 1994)
CREAM	Cognitive Reliability and Error Analysis Method (Hollnagel and Embrey, 1994)
CAMEO/TAT	Cognitive Action Modelling of Erring Operator/Task Analysis Tool (Fujita *et al.* 1994)
CREWPRO†	CREW PRObem solving simulation (Shen *et al.* 1994)
SRS-HRA	Savannah River Site HRA (Vail *et al.* 1994)
EOCA	Error of Commission Analysis (Kirwan *et al.* 1995)
SYBORG	System for the Behaviour of the Operating Group (Takano, *et al.* 1996)
SEAMAID	Simulation-based Evaluation and Analysis support system for MAn-machine Interface Design (Nakagawa *et al.* 1996)
ATHEANA	A Technique for Human Error ANAlysis (Cooper *et al.* 1996)

Note: Acronyms marked with an '*' were reviewed in Kirwan (1992a, b), those marked † are this author's acronyms for the techniques, since the authors did not supply one in the original reference reviewed.

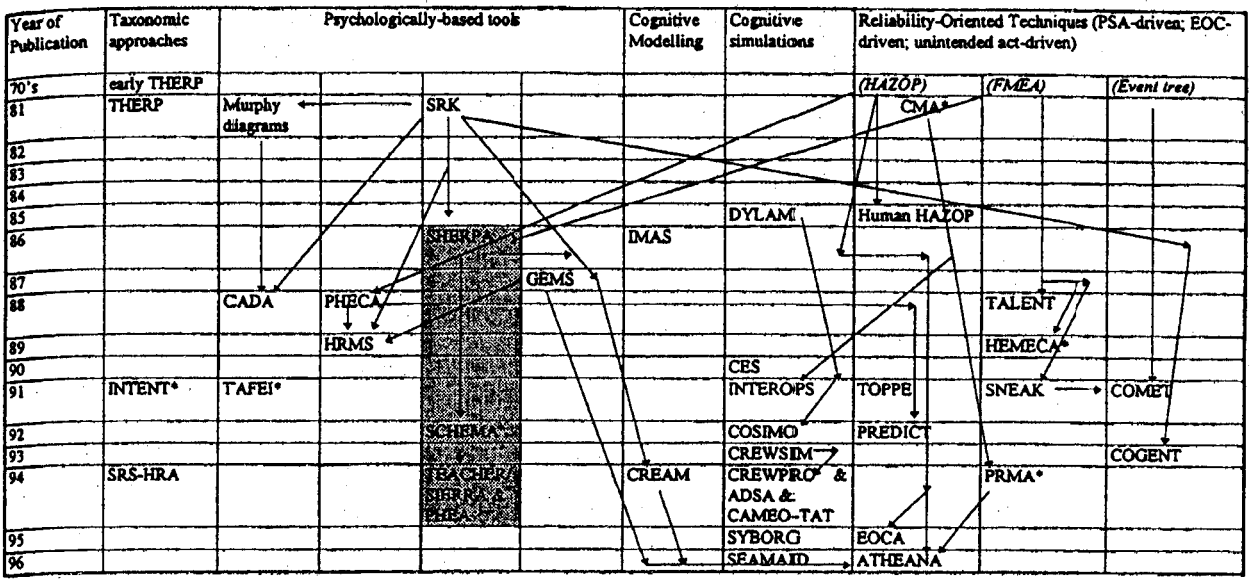
The techniques themselves are described in brief below, within these categories, and trends in their development are noted, prior to their evaluation. However, twelve of the techniques above (those marked with a single asterisk in *Table 1*) were reviewed in Kirwan (1992a, b), and are described in Kirwan (1992a). These techniques are not therefore re-described here.

(i) Taxonomic approaches

Taxonomic approaches have been used for over three decades to identify human errors in risk and reliability assessments, the most used technique being THERP's error taxonomy (e.g. *error of omission*; *wrong timing*; *wrong sequence*; *wrong action*; etc.). Such techniques are either generic (like THERP), or specific and related to a specific industry (such as SRS-HRA), or a specific error type such as rule violations (e.g. INTENT).

The Savannah River Site HRA (SRS-HRA: Vail *et al.*, 1994) method is a data-based approach based on data collected from four existing SRS databases (based on incidents, logs, etc.): fuel processing; fuel fabrication; waste management; and reactors. The approach is contextual and taxonomy based, and in this respect is similar to the recent JHEDI approach (Kirwan, 1990, linked to HRMS), another nuclear fuel processing facility HRA method. The list of errors that could be used in SRS risk analysis event and fault trees was designed to be

Table 2 Approximate evolution relationships of HEI techniques



Key: Italics – reliability techniques; shaded column area (SHERPA etc.) – a major distinct 'family' of techniques with logical and intended 'descendants'.
Note 1. Where there are more than one technique in a box, arrows to one technique are for that technique only, i.e. not all techniques in a box. 2. * These techniques could appear in more than one column – e.g. INTENT could be applied as a HAZOP-based approach; TAFEI could be considered to be reliability – oriented (under FMEA) approaches, as they break down the tasks and look for failure opportunities based on system 'affordances'; PRMA could be considered a crude cognitive modelling approach, as could CMA; and HEMECA could be considered a psychology based technique as it uses PSF (but is very reliability-oriented).

comprehensive, and the listing was therefore based on a mixture of incident databases, reviews of prior risk assessments, and judgement. Examples of error modes are: failure to notice/respond to an alarm; miscalibration; failure to restore following maintenance; failure of administrative control; and laboratory analysis error.

The method INTENT [not an acronym] (Gertman, 1991) is aimed at enabling the incorporation of decision-based errors into PSA, i.e. errors involving mistaken intentions, which appears to include cognitive errors and rule violations, as well as EOCs. Four categories of error of intention were identified: action consequence (e.g. *tolerating an out-of-range situation with potentially major consequences*); crew response set (e.g. *incorrect interpretation of symptoms*); attitudes leading to circumvention (e.g. *violate procedure and reconfigure equipment*); and resource dependencies (e.g. *inadequate communication results in improper actions*). A set of 20 errors of intention (and associated PSF) were derived, and quantified using seven experts.

The taxonomic approach can therefore be either (behaviourally) generic, contextual, or focusing on a subset of error types. In all cases the approach is essentially experience based: incident experience tempered with assessor experience. This means it has a certain degree of context validity and is low in resources usage, depending on the assessor and his/her understanding and experience of the context being analysed. However, if novel contexts are being assessed, the technique/assessor may fall short in identifying novel human interactions and errors arising. This is why other techniques try to be more model based—they are more resource intensive, but in theory are capable of being more comprehensive. Whether they are in practice, compared to a purely taxonomic approach, will depend upon analyst skill and understanding of the task and its underlying error forms and error causative factors.

(ii) Psychologically-based tools

Task Analysis For Error Identification (TAFEI: Baber and Stanton, 1991; Stanton, 1994) is a task analysis method based on State Space Diagrams, describing user interactions with equipment in terms of transition (input–output) boxes (non-Markovian: qualitative in nature). For a particular task the network of transition boxes is developed, and then examined to determine what illegal transitions could take place, such as skipping over task elements, sequence errors, etc., though in theory EOCs could be developed from such networks (e.g. as with SNEAK–see below). The philosophy underlying TAFEI is that of '*affordances*' (proposed by Gibson, 1979), i.e. that error is a function of what the designed system affords (allows) the operator to do, as opposed to how the system was designed to operate. TAFEI could be developed to determine equipment design robustness against EOCs, and the potential for EOCs and Rule violations, particularly with respect to mode errors in aviation, for example. It is however a resource-intensive approach, and the transition matrix and State Space Diagrams may rapidly become unwieldy for even moderately complex systems.

A number of methods have evolved from the original SHERPA method (Embrey, 1986), and have generally followed its flowchart format, such as the Systematic Critical Human Error Management Approach (SCHEMA: Livingston *et al.*, 1992). The most recent variant however is the Predictive Human Error Analysis

(PHEA) technique, within the TEACHER-SIERRA framework (see below). This technique comprises an error checklist (External Error Modes or EEMs) and is a simplified version of the earlier SHERPA technique developed by the same author. The technique similarly requires a Hierarchical Task Analysis (HTA) and focuses on particular task types depending on the industry concerned (e.g. planning; actions; checking; and information transmission/communication for the chemical industry). Within each task type, there are a set of EEMs (e.g. *action mistimed*; *check mistimed*; *information communicated at wrong time*). The output is a SHERPA-type table (task step; error type; description; consequences; recovery; remedy). There appears to be no consideration of PSF/PEMs within the PHEA system, these being left to SIERRA, the other main technique within the TEACHER framework.

The Technique for Evaluating and Assessing the Contribution of Human Error to Risk (TEACHER) framework (Embrey, 1993) appears to be an alternative HRA framework more aimed at lower consequence accidents than PSA traditionally aims at. It has a number of components. The first is the Systems Induced Error Approach (SIERRA). This states that humans have basic error tendencies which are influenced by certain factors (called Performance Influencing Factors, or PIF). TEACHER focuses on defining a task inventory, then determining the prioritisation of critical tasks according to their risk potential (via a pre-defined structured weighting system), leading to a rating on a risk exposure index for each task. This whole process would seem to require significant judgement. Following the screening analysis a HTA and PHEA analysis are carried out, following which, those errors with significant consequence potential are analysed with respect to a set of PIF audit questions, to develop remedies for the error. Each PIF audit question allows the analyst to rate the task according to, e.g., the extent to which procedures are defined and developed by using task analysis, on a seven-point semantic differential, anchored at each end-point. Risk reduction is then determined by the analyst. This flowchart-based set of approaches is effectively an extension of the taxonomic one, but lends more structure to the error identification process, making it more reliable. The flowcharts also have more basis in theory (e.g. SRK theory).

The psychology based tools (including Murphy Diagrams, PHECA, CADA, and HRMS, all of which are apparently not in current use or else used rarely, and GEMS and SRK, which are rarely used as tools on their own, and act more as a basis for other technique development) occupy a difficult niche in human error identification. This is because they attempt to bring generalised psychological theories or models into the rich context of a complex industrial work environment. Techniques such as SHERPA (and PHEA), SRK, and TAFEI will be able to identify a good number of error forms from their basic approaches, but may founder in truly complex task environments, or simply become unwieldy or too resource intensive. The development to watch for the future could well be CREAM, defined below, since it offers the potential of combining a pragmatic approach such as SHERPA with more useful cognitive psychological back-up, leading, in theory at least, to more accurate and insightful error identification and generation of error reduction measures.

(iii) Cognitive modelling

The Cognitive Reliability and Error Analysis Method (CREAM: Hollnagel and Embrey, 1994) attempts to bring cognitive psychology/science into the HEI arena, i.e. CREAM is aimed at being a more theoretically valid approach. It is a compound of SHERPA, SRK, and the COCOM (COgnitive COntrol Model: Hollnagel, 1993) approach, the latter stating that HEA can only occur accurately based on a detailed and psychologically valid modelling of the context of the task in its environment. The approach can be applied retrospectively or prospectively, although it does not yet appear to have been used prospectively, and the authors themselves note that further development is required. It also has an overall taxonomy for human error analysis. The 'meat' of CREAM discussed in the paper is the Action-Error-Analysis Matrix. This shows relationships between 'causes' and 'effects', in both cases being a non-mutually-exclusive mixture of error mechanisms and performance shaping factors and some external error modes, occurring on both axes. The system is still under development (Hollnagel, 1996).

This technique and the other in this column in Table 2 (IMAS: not currently in use) are different from the previous technique category in that they trying to model cognitive behaviour. Whilst other techniques (e.g. SRK; HRMS; CADA; etc.) may identify knowledge-based/cognitive errors (e.g. misdiagnosis), such techniques do so only in a superficial way, not exploring the cognitive behaviour itself. Instead they focus on the overall cognitive error types that can occur, given a basic task analysis that indicates that some cognitive effort is required, and therefore cognitive failures may occur. From the Ergonomics and cognitive psychology perspective, techniques in this category are desirable, as they should be more accurate and diagnostic due to a more theoretical framework, and should be more capable of dealing with cognitive aspects of complex tasks. More development is therefore clearly needed in this category, and could be linked to cognitive task analysis approaches.

(iv) Cognitive simulations

There are a number of cognitive simulation approaches. These, each, have the ambitious aim of developing a computerised simulation of operator performance based on some underlying model of performance (e.g. information processing; symbolic processing theory; etc.). As a consequence such models can predict non-performance or errors in tasks. Most of these methods are aimed at knowledge-based (or at least rule-based) behaviour, which is one of the more difficult error prediction domains. The simulation itself needs a software architecture (e.g. blackboard architecture) to act as an environment to 'house' the model of performance. Furthermore, several of the models then additionally utilise sub-models (e.g. fuzzy set theory; Markov modelling; etc.) to account for certain specific types of behaviour (e.g. decision-making under uncertainty). These simulations usually require years to develop and represent significant programming effort. They also tend to require significant subject matter expertise (SME) or performance data to determine how operators would react in various conditions. Nevertheless, once computerised, the system will predict performance and errors for

a range of scenarios, and as such can be validated against real human performance. Five of the ten cognitive simulation approaches listed in *Table 1* are briefly reviewed below[†] (these five give a good flavour of the various techniques' functionality and approach).

(1) The Integrated Reactor Operator System (INTEROPS: Schryver, 1991; Schryver and Knee, 1991) is a cognitive performance simulation developed at Oak Ridge National Laboratories, which uses the SAINT (System Analysis of Integrated Network of Tasks) simulation methodology. INTEROPS has three independent models: a NPP model; a network model of operator tasks; and a knowledge base, the operator model being distributed between the latter two. The model is a single operator model. It diagnoses by observance of plant parameters, and subsequent hypothesis generation and testing of the hypothesis. The INTEROPS model allows the following to be simulated: forgetting, tunnel-vision; confirmation bias; and mistakes. The approach uses Markovian modelling to allow opportunistic monitoring of plant parameters. Stress is 'accumulated' as the emergency proceeds from detection of a disturbance, and peaks at diagnosis achievement, and then decreases monotonically (assuming diagnosis is correct). The probability of ignoring evidence (tunnel vision) increases monotonically with time pressure, and confirmation bias is modelled by not testing for fault paths not identified in the initial fault proposition phase. The model also simulates various errors such as reduced depth of analysis due to time pressure, task switching, and inherent deficiencies in the knowledge base. PSF are also used in INTEROPS, namely training, time pressure, sleep deprivation, and tension. Cognitive workload is also modelled, in terms of the contemporary information processing theory of concurrent task management. Also, INTEROPS can utilise a confusion matrix approach to make diagnostic choices.

(2) The Cognitive Simulation Model (COSIMO: Cacciabue *et al.*, 1992) is a parallel to CES (see earlier) in that it is a simulation of the human operator and his/her thought processes, using a computerised blackboard architecture. The simulated operator comprises a set Rule-Based Frames (RBFs), which are a set of properties and attributes associated with particular incident scenarios (including the incident's relative experienced frequency), and Knowledge Based Frames (KBFs) which are 'packets' of process knowledge (e.g. physics principles, etc.) and heuristics/rules of thumb. When diagnosing, each scenario and its associated attributes are contrasted to 'similarity-match' to the symptom set being displayed to the 'operator', and the simulated operator will either determine unequivocally which scenario matches the symptoms, or, if there is ambiguity, will 'frequency-gamble'. The COSIMO system however takes into account the diagnosticity of symptoms. Once hypotheses are formulated, they are evaluated according to a confidence threshold, and may be accepted or rejected. This threshold can be dynamically modelled during the COSIMO run. This feature could be used to model, for example, reluctance effects or the effects of stress or over-confidence on decision-making and diagnosis. By altering other features of the mechanics of the system, a range of cognitive errors can be 'caused', such as cognitive lock-up, or cognitive collapse.

(3) The CREW SIMulation model (CREWSIM: Dang *et al.*, 1993; Dang and Siu, 1994) is a simulation model that models the response of an operating team in a dynamically evolving scenario, and has been particularly developed to date to focus on a particular nuclear power plant scenario. The model simulates operator interactions within a three person crew, as well as the cognitive processes of the crew members, and the crew-plant dynamic interaction. Although the model has a knowledge base as other simulations do (e.g. COSIMO and CES), CREWSIM differs by using a set of prioritised lists that reflect the priorities of different concerns. Production rules can also be used for diagnosis, and memorised tasks are represented as scripts within the system. Scripts are therefore used to simulate rule-based behaviour. CREWSIM has some other interesting aspects. Firstly attentional resources control is simulated, such that diagnosis will be suspended while the operator is communicating or carrying out some other task. Secondly, the model's usage focuses particularly on transitions between procedures, and hence is looking in particular for premature, delayed, and inappropriate transfer within the emergency procedures system. Thirdly, several error mechanisms are treated by the model: memory lapse; jumping to conclusions; communication failures; incorrect rules; and improper prioritisation. Communication errors may be modelled as a function of self-confidence and confidence in other members' abilities, and message transmission/receipt can be affected by stress (but garbled messages are not represented). Incorrect rules can be modelled by 'tampering' with (i.e. degrading) the knowledge base, as with other simulation models, but also can be achieved by influencing rule strength in CREWSIM.

(4) The Cognitive and Action Modelling of Erring Operator/Task Analysis Tool (CAMEO/TAT: Fujita *et al.*, 1994) is a simulation approach acting as a task analysis tool, primarily to evaluating task design (e.g. ease of use of procedures), but also for potential use in HRA. It allows designers to ensure that operators can carry out tasks. PSF used in the approach include task load, complexity, time pressure, opportunistic change of task order, multiple task environments, negative feedback from previously made decisions or actions, operator's policies and traits, etc. The CAMEO/TAT system is primarily information processing containing psychological (simulated) 'modules' such as working memory, long-term memory, etc. Errors are modelled mainly as a function of insufficient resources. Designers then manipulate the design of the system until no more erroneous tendencies occur. The system is a single operator simulation. The total amount of resources may vary as a function of stress. Furthermore, task switching mechanisms exist to determine the next task to carry out, and these mechanisms can be opportunistic, random, or linked to likely preconceptions and dispositions of the operator. In this way, pattern matching strategies (and associated errors such as similarity matching) can be modelled. One interesting feature is the individual differences implied as variable in the simulation, so that, for example, the simulation could model an operator who would switch diagnoses more often, etc. (possibly reacting too fast leading to premature acts based on

[†]CES has already been presented in Kirwan (1992a), and DYLAM has been partly been superseded by COSIMO. ADSA and CREWPRO build on CREWSIM, and SEAMAID has a functionality similar to CAMEO-TAT. For more information on these techniques see also Kirwan (1996) and Kirwan and Hollnagel (1998).

insufficient evidence). This approach is relatively rare in HEI,[1] where more usually either an 'average' operator is considered, or a conservatively worse than average one is conceptualised.

(5) The System for the Behaviour of the Operating Group SYBORG (Takano *et al.*, 1995; Hasegawa and Yoshimura, 1996) is a recent cognitive simulation approach which is the first to try to deal with emotional aspects of performance. It aims to predict what emotions personnel will experience when dealing with difficult nuclear power plant events, and aims to determine how these emotions will affect attention, thought, action, and utterances. The emotions considered include fear, anxiety, tension, surprise, etc. There is ongoing work to determine how emotions interact with each other and with error forms, e.g. their research suggests that 'indecision' is linked to *discouragement*, *dislike* and *irritability*, but not if *tension* and *satisfaction* are activated (Hasegawa and Yoshimura, 1996). These complex interactions and inter-relationships are based on empirical observations of actual subjects performing in simulator experiments. SYBORG is possibly the first approach that, in the future, may be able to identify idiosyncratic errors, or errors caused by extreme stress in a situation.

The cognitive simulations therefore differ in terms of their underlying theory, sub-models, and objectives—some are aimed at error analysis, others at design evaluation (see also Kirwan, 1996; Kirwan and Hollnagel, 1998). These approaches clearly represent the most sophisticated end of the human error identification domain. The main concern with these systems, however, is that they rarely seem to reach the stage of maturity where they can be used in actual evaluations (PSAs), instead often remaining as research tools. It would be more useful to see some of these approaches applied in real PSAs, to gauge their true utility.

(v) Reliability-oriented techniques

A number of techniques have adopted reliability engineering-style approaches to human error identification. The two main reliability engineering approaches are Failure Mode and Effects Analysis (FMEA), a structured single-assessor detailed analytical approach, and Hazard and Operability (HAZOP), a semi-structured, group approach. The former approach is often seen as deductive in nature, and to be looking at failure possibilities both systematically and in great detail. The latter HAZOP approach is seen to be more inductive and hence incisive, and verging on brainstorming at times. Whilst reliability of approach is likely to be higher with FMEA, depth of insight is often higher with HAZOP, particularly when dealing with less easily structured problems or novel systems. With HAZOP, resources can be high. Also, two counter-posing adages are worth remembering with HAZOP in particular: first, that several heads are better than one; and second, garbage in, garbage out—the quality of the subject matter experts in a HAZOP will determine its utility and efficacy.

A third style of reliability engineering approach that has been adopted for developing HEI techniques is the event tree technique, in which the progression of possible events, in time and commencing from a common starting point, is modelled. Event trees have been used for some time to model the different actions operators can take given a procedure (as in THERP, for example), so it is not surprising that they are used as a platform for HEI considerations.

This category of techniques is therefore divided into three sub-categories: HAZOP or group-based techniques; FMEA-type techniques; and event tree-based techniques.

HAZOP and group-based techniques

The Procedure to Review and Evaluate Dependency In Complex Technologies (PREDICT) method has been proposed by Williams and Munley (1992), but to the author's knowledge has not been formally applied. It is targeted at the relatively unpredictable or bizarre event sequences which characterise events from TMI to the Herald of Free Enterprise, in that such events are incredible or not predictable until accidents give us 20:20 hindsight. The method utilises a group to identify errors, and is thus HAZOP-based, with Kletz's (1974) and Whalley's (1988) keyword systems (e.g. 'no action'; 'action repeated'; etc.), followed by three categories of assumption-testing keywords (low, medium and high severity challenge: e.g. 'confirmed by'; 'not recalled because'; and 'defeated by'). The technique essentially allows the analyst to test the assumptions underpinning the design and safety cases for plants. The paper also mentions a facility to insert a keyword randomly to enable the analyst to consider more 'lateral' possible causal connections. Exactly how effective or resource-intensive this method would be in practice is difficult to say. It is also not clear how easy it is to isolate all the key assumptions underpinning design/safety cases. The approach is however unusual, and takes the open-ended and open-minded philosophy of HAZOP to a more extreme position. It currently occupies a unique potential niche in identifying errors of commission, idiosyncratic errors, and rule violations.

Error of Commission Analysis (EOCA) is a HAZOP-based approach whereby experienced operators consider procedures in detail, and what actions could occur other than those desired (Kirwan, 1994; Kirwan *et al.*, 1995, 1996). Particular task formats, error mode keywords, and PSF are utilised to structure the assessment process and to prompt the assessors. Identified significant errors are then utilised in the PSA fault and/or event trees. This approach has only been used once, albeit successfully, in a real PSA.

A technique for Human Error Analysis (ATHEANA: Cooper *et al.*, 1996) has been developed relatively recently to analyse operational experience and understand the contextual causes of errors, and then to identify significant errors not typically included in PSAs for nuclear power plants. These errors may well be errors of commission, and their identification relies heavily on an understanding of the context surrounding the performance of tasks. These contextual factors amount to plant events and anomalies (e.g. incorrect readings from indications etc.) and PSF. The ATHEANA process in brief requires that key human failure events and associated procedures etc. are identified from the PSA (e.g. operator fails to start pumps in an emergency), and unsafe acts (e.g.

[1]The SEAMAID system (Nakagawa *et al.*, 1996) has a similar overall approach and purpose to CAMEO-TAT but the underlying model has operator knowledge represented in Petri Net form.

running equipment is turned off) are then identified that could affect or cause these events. Associated error-forcing conditions (e.g. badly structured procedures; misleading indications) are then identified that could explain why such unsafe acts could occur. The important point is that these forcing conditions are based on the system being assessed, i.e. the real context that is the focus of the assessment.

Currently the identification of events, unsafe acts, and error-forcing conditions and PSF are via review of operational experience and expert judgement by the analysis team (engineers and Human Reliability practitioners). It is the intention of the authors to produce guidance material on the technical basis of the model (human performance models and how errors are caused, based on theory and operational experience). Such material could reduce the reliance on expert judgement and increase the auditability of the technique.

The Team Operations Performance and Procedure Evaluation (TOPPE) technique (Beith *et al.*, 1991) is a procedure validation and team performance evaluation technique. It uses judges to evaluate team performance when carrying out emergency procedures. It is therefore not designed as a HEI tool. However, it can identify procedural errors (omissions, wrong procedural transitions, etc.), and team leadership or co-ordination problems. As such, an approach could be developed to determine credible procedural and co-ordination errors of these types, based on observation of emergency exercises which all NPP utilities are required to carry out.

FMEA-based techniques

Human Error Mode Effect and Criticality Analysis (HEMECA: Whittingham and Reed, 1989), is a FMECA-type approach to HEA. It uses a HTA followed by error identification and error reduction. The details of the error mode identification are not given in the paper, but the PSF used by the analyst are primarily man–machine interface related, e.g. workplace layout, information presentation, etc. The paper notes that typically an FMEA approach identifies many errors, primarily through detailed consideration of these PSF in the context of the system design, in relation to the capabilities and limitations of the operator, based on Ergonomics knowledge. Only those errors that are considered to be probable within the lifetime of the plant are considered further, i.e. all others are screened out (an inappropriate criterion for HRA/PSA). Error reduction is a fundamental objective of this approach, based on the PSF and detailed analysis of the interface, and each error identified is prioritised for reduction via a criticality rating, leading to the potential to carry out cost–benefit analysis.

The Task Analysis-Linked EvaluatioN Technique (TALENT: Ryan, 1988) is an assessment framework which also contains a strong task analysis bias, utilising task analysis (presumably), some form of linear HTA or sequential task analysis, timeline analysis, and link analysis for each task sequence (see Kirwan and Ainsworth, (1992), for a description of these task analysis techniques). Then, tasks are identified for inclusion in the fault and event trees, through a collaborative effort between the behavioural scientists and the safety assessors. PSF are then identified for each task, and then the tasks are quantified using either THERP or SLIM. TALENT was apparently applied in a large European PSA/HRA exercise, and for an evaluation of the US Peach Bottom nuclear power plant. As far as the author is aware, TALENT has not been used substantially recently.

SNEAK (Circuit) Analysis was originally developed to look at unintended connections in wiring systems, and was then adapted considerably to consider errors of commission (EOCs) in HRA by Hahn and de Vries (1991). Sneak Analysis starts with the development of a stepwise flowchart of the task sequence. Clue application is the next component of Sneak Analysis, carried out using the computerised system the authors have developed. If for example, information must be sought by the operators, a number of questions are asked of the analyst in flowchart form (can the operator misread/mishear the information?; etc.). A number of the questions will require a relatively detailed human factors analysis of the installation if they are to be answered. For each question, there is back-up information expanding on what constitutes an acceptable system configuration in human factors terms. Sneak paths are then identified by considering the logical possibilities for flows in the system, i.e. based on the actual system configuration and possibilities. Barriers that are present must be considered at this point. For example, it may be possible to open the wrong valves, but there may be a key-access system on plant (controlled by the Shift Supervisor) which will be a possible source of prevention of this error.

The human engineering deficiencies identified in the middle phase of Sneak Analysis appear to be used to influence the probability of the EOC, as well as helping to identify them. The approach appears to be highly resource intensive, but particularly apt for identifying 'Right-action-wrong-object' [a SHERPA EEM] and sequence (e.g. sneak timing) errors. However, these are a subset of EOCs, and given the high resources investment to identify them using SNEAK, it must be questioned as to whether there are not quicker, more incisive means of determining EOCs.

The Procedure Response Matrix Approach (PRMA) (this title given by this author) has been recently proposed by Parry (1994) for identifying errors of commission, which in his terms are more closely linked to cognitive errors (global and local misdiagnoses), and slip-based EOCs during emergencies. The approach has strong affinities with the Fault Symptom Matrix Approach (FSMA—linked to the Confusion Matrix Approach, CMA: Potash *et al.*, 1981; Kirwan, 1994), which has faults on one axis of its matrix and symptoms on the other one. PRMA to some extent represents a more sophisticated and detailed investigation than the FSMA, though one which is more resource intensive. The approach has several major stages: develop a PRM for all initiating events that produce significantly different plant responses; for each PRM review the decision points in the procedural pathway; identify potential incorrect decisions resulting from misinterpretation or failure of the plant to provide the appropriate information, or due to a procedural omission (lapse). The critical indications can be reviewed to see if there are redundant and diverse indications of important signals, and this is summarised in a Plant Information Matrix (PIM: a form of FSMA), listing the critical parameters and their respective levels of redundancy etc. Recovery is considered as a function of procedural direction, unless there is a potential for mindset due to, for example, a training bias.

Additionally, single instrument failures are postulated, and their likely effects on performance considered in the light of the PIM (e.g. via redundancy and/or diversity in instrumentation). Misleading indications are also investigated (e.g. due to effects of secondary failures or cascade effects etc.). Lastly, a number of PSF are considered, namely workload, perception of time urgency, perceived reliability of instrumentation, remoteness of instrumentation, clarity of procedures, and training under/over-emphasis of scenarios. Assessment of these factors is used to determine the likelihood of ignoring (dis-) confirming information. The technique is useful for considering how system status indications and procedures will affect performance in abnormal or emergency events, such as a nuclear power plant emergency scenario requiring diagnosis and recovery actions using emergency procedures. As such it can be used to evaluate alarm system design adequacy, for example.

Event tree-based techniques

Modified event trees, called COMmission Event Trees (COMETs: Blackman, 1991) deal with errors of commission and cascading errors whose source is either erroneous intention or a latent error. COMETs are developed, e.g. using SNEAK (see above), and are basically event trees, their results feeding into fault trees. The main significance of this approach appears to be as a means of integrating errors of commission into PSA and quantifying them. It does not help too much in terms of actually identifying errors of commission.

The COGnitive EveNt Tree System (COGENT: Gertman, 1993) is another extension (as with COMET) of the THERP event tree modelling system, this time dealing particularly with cognitive errors, although the approach appears to deal with other errors as well (S,R, and possibly EOCs). The aim is to bring current more cognitively based approaches into the HEI process, i.e. Rasmussen's and Reason's taxonomies. This has led to a hybrid taxonomy with terms such as 'Skill-based slip', rule-based lapse, and knowledge-based lapses or mistakes (there is a distinction between 'simple mistakes' and 'mistakes'). The approach thereafter is for the analyst to develop cognitive event trees. It requires significant analytical judgement. At present, it appears to be a relatively simple step forward in modelling (representation), rather than in HEI.

Discussion of approaches

As well as classifying the approaches in terms of their general theoretical direction, the various techniques can also be classified in terms of their analytic method, i.e. how they are used in practice, which enables a clearer focus on the likely efficacy of the various tools. The techniques appear to fall into a number of such categories (which are not mutually exclusive), as briefly discussed below.

(i) checklist-based approaches (e.g. SRS-HRA, THERP; INTENT; GEMS)

These approaches are relatively easy to use whether by novice or more experienced practitioner (the results may obviously differ). In reality although appearing simple and straightforward, they rely on the skill of the assessor, and the degree to which the assessor understands the task being assessed. These are really prompts. One drawback is that current checklists tend not to be very cognitive in nature.

(ii) flowchart-based approaches (e.g. SHERPA and its derivatives)

Flowcharts offer considerable structure to the assessor, which is one of the reasons that SHERPA, for example, has retained its popularity. These are of special utility for novices, who will not have sufficient experience to know when to apply certain error identification keywords. Flowchart-based tools are likely to be highly auditable, and to lead to more reliable or consistent assessments between different assessors. They are therefore more appealing in areas where other comparable assessment approaches have a high degree of quality assurance and come under a significant amount of scrutiny. Flowchart-based techniques are also more likely to come from theoretical models. The drawback here is that, as with many other forms of HEI tool, models which lend themselves to flowchart derivatives only tend to deal with more straightforward types of behaviour (e.g. skill and rule-based behaviour). Flowcharts for more cognitive behaviours, such as decision-making and problem-solving, which often can have most impact on the risk of an installation, are few and far between.

(iii) Group-based approaches (e.g. HAZOP; PREDICT; EOCA; CMA)

Group-based approaches such as HAZOP and its 'relatives' clearly have a place in error identification, and are likely to be used for some time. HEI is obviously a difficult area, and one in which psychological foundations and the resulting scientific methodology available is lacking. In such an instance, it is sensible to attempt to harness the implicit knowledge of experienced personnel, and to use group processes to elicit less obvious error forms. These group processes facilitate inductive thought. This is particularly important when predicting events for novel systems, where such events may not have happened anywhere yet. In such cases, more formalised methods may find it difficult to extrapolate from known to unknown contexts. This is an area where the human can make such extrapolations, supported by documentation, checklists, experience of other contexts, and support of other group members. The disadvantages with group-based approaches are firstly that their reliability cannot be guaranteed, and secondly their cost. These disadvantages must be balanced against the potential benefits of comprehensiveness and insight, particularly for novel system designs.

(iii) Cognitive psychological approaches (e.g. GEMS, SRK, CREAM)

These approaches are potentially of most interest to psychologists and others who want to predict the more sophisticated error forms associated with misconceptions, misdiagnoses, etc. They attempt to explore the error forms arising from 'higher-level' cognitive

behaviours. SRK in particular has been a landmark effort, and has created a bridge between reliability engineering and risk assessment on the one hand, and psychology and Human Factors and Ergonomics on the other. Yet it is limited, and new approaches are required, whether building on systems such as GEMS, or more novel hybrids such as the prototype CREAM technique which is still under development.

(iv) Representation techniques (which ease integration into PSAS: e.g. COMET, COGENT)

These techniques map very well into PSA approaches, as they are designed to do so. They tend to be useful in maintaining the direction of the study, by highlighting what the risk assessment needs from the HRA, and they never lose sight of identifying error forms which can be integrated into the PSA. However, they tend to require other methods to identify errors at a lower level which will feed into the event tree representations. These approaches are very good at providing the structure for error impact to be represented in the PSA, but less so at providing the content of that error impact. Once again, this is less a weakness of these techniques themselves, and more a result of the lack of detailed cognitive models of error causation.

(v) Cognitive simulations (e.g. CES, COSIMO, etc.)

A notable trend is the development of a range of computer simulation-based models: whether these are based on an information-processing theoretic approach, or a symbolic processing framework. Such simulation approaches invariably originate from the cognitive science domain rather than Human Factors. These simulations model the operator's thought processes, and offer potentially powerful ways of determining how human operators will respond in emergency scenarios, typically in complex environments such as nuclear power plants.

Cognitive simulations are without doubt the most ambitious and resource-intensive approaches, at least in terms of their development requirements. Their strength is that once completed, they can predict human performance with each simulation 'run', and do not require analytical judgement or reliance on groups of experts. Some of them have shown evidence of validity (e.g. CES) in predicting actual operator performance. They also offer means to identify the most difficult cognitive error forms (e.g. misdiagnosis, etc.), and to suggest when and where such errors might arise during a complex procedure execution.

There are two main disadvantages however. Firstly, the simulations are only as good as the cognitive models underlying them, and as already noted, such models are still not fully mature. Secondly, such cognitive simulations are very resource-intensive to develop, and a number of models have been partly developed only to lose funding and fail to reach the stage of actually being used for risk assessment purposes (Kirwan and Hollnagel, 1998).

(vi) Task analysis linked techniques (e.g. TALENT; PRMA; CAMEO-TAT; SHERPA)

This style of analysis focuses on the detail of what the operator must do, and therefore relies on a detailed task analytical approach. This approach is very relevant to Human Factors. It is also more documentable than some of the other approaches, allowing more of an 'audit trail' for the assessment. The more recent approaches, such as PRMA and to an extent ATHEANA, are also extending the analysis of the *context* in which the operator operates at the time, which could influence the operator and lead to some the more complex error forms (EOCs; cognitive errors). This shift towards analysing and representing the context is entirely in line with the Human Factors (and psychology) philosophy of approach. It also means that error reduction measures will be more easily identifiable and more relevant to the system being investigated.

(vii) Affordance-based techniques (e.g. TAFEI; SNEAK; PREDICT)

These approaches are interesting because they represent a systems and hybrid approach to error identification. Essentially, they start from the viewpoint that anything that can in theory happen, can happen (such events are 'afforded' by the system architecture and its operational environment). What they then try to do, given such a very large 'problem space', is consider why such human actions could happen, establishing reasons for such actions. This approach differs from most others in that it looks for possible actions and then tries to establish potential intentions (erroneous or not) for such actions. Most other approaches start from intentional action and then look for potential deviations or wrong conditions for such intentions. The affordance approach is particularly relevant to errors of commission and rule violations, the areas least defended in current high technology high risk systems.

(viii) Error of Commission (EOC) identification techniques (e.g. SNEAK; EOCA; ATHEANA)

Error of commission analytic techniques represent another significant trend, with a range of techniques and approaches being suggested for a range of industries. Such errors are very frequently due to a subtle interplay between poor ergonomics design aspects, and can sometimes prove disastrous. Some techniques for identifying such errors are table-top simulation approaches, others relying on databases of ergonomics guidance to help identify the potential for slips (unintended acts) leading to errors of commission. SNEAK relies on an Ergonomics database, and is philosophically interesting in that it is attempting to derive EOCs at a very detailed level, and is considering essentially previously unconsidered connections between events, actions, and system states. It is also potentially a very useful tool, since ergonomics design deficiencies are used to determine where and when sneak paths will occur. It is however a highly resource-intensive technique, and cannot be applied until the operational stage of the system. ATHEANA, the latest tool to tackle this area, is also interesting as it can be seen as trying to give credence to a difficult assessment area by drawing from operational experience, backed up by theory. Such an approach may overcome the residual scepticism from the engineering world concerning the real threat EOCs pose to installations.

(ix) Crew interactions and communications (e.g. CREWSIM; CREWPRO; TOPPE)

Another noticeable recent trend is the attempt to model operating crew interactions, mainly within simulation methods, in terms of modelling communication errors and team members' confidence in their colleagues abilities and judgements. This modelling level represents the most psychologically ambitious so far, and the most realistic in terms of actual crew co-ordination tasks in a nuclear power emergency. CREWSIM is currently the only simulation model that models (albeit crudely) team interactions in a dynamic event. Three members of the crew are simulated. CREWSIM models attention resources control, such that diagnostic activities will be suspended while other activities must be attended to. Again, as with ADSA, this represents an increase in modelling the dynamic aspects of the evolution of a scenario. CREWSIM also allows the consideration of different goal priorities within the operator, and in particular focuses on transitions from the basic generic emergency procedure in a certain reactor system design to other procedures (e.g. to Steam Generator Tube Rupture). Communication and confidence in other crew members are also intended to be modellable in an extension to CREWSIM called CREWPRO. These represent ambitious but significant enhancements of the external validity or realism of modelling (though of course such modelling requires validation). As yet, however, there seems to be no definitive communication error taxonomy or model upon which to base such errors, and hence approaches such as TOPPE could be of benefit if adapted towards such a purpose.

The above observations represent overall characteristics of the techniques, and trends in their development. The next section in this series evaluates each technique against an exhaustive and formal set of criteria, to determine which techniques are particularly useful, and to determine where new developments are required.

Criteria for evaluation of HEI techniques

The criteria set for evaluating the techniques are an expanded set from the Kirwan (1992b) review of twelve HEI techniques. In that review, a criteria set was developed as described below.

Original criteria set (Kirwan 1992a, b)

Two related criteria sets were used in the above-mentioned review. The first was useful for HEI comparative validation exercises, and the second for more qualitative evaluations. Since formal empirical validation is outside of the scope of this project, the latter criteria set is the one that is used in this review.

(i) *Comprehensiveness of human behaviour*: the degree to which the technique addresses skill, rule, and knowledge-based behaviour, rule violations, and errors of commission etc. (abbreviated to S, R, K, RVa, and EOC, respectively) (see Section 2 in Kirwan, 1996).

(ii) *Consistency* (in terms of the degree to which the technique is *structured*, and so more likely to yield consistency of results, versus a technique which is open-ended, in which case the results are likely to be highly assessor dependent). Techniques are rated as low (meaning a relatively open-ended technique), moderate (meaning that the assessor has flexibility within a detailed framework), or high (meaning that the tool is highly structured and likely to lead different assessors down the same error identification routes, given the same information and assumptions).

(iii) *Theoretical validity* 1: whether the technique is based on a *model* of human performance. Techniques are rated as low (indicating a simple classification-based system), moderate (indicating that the technique makes reference to a model of human performance), or high (meaning that the tool is an embodiment/interpretation of a model of human performance).

(iv) *Theoretical validity* 2: whether the technique simply assesses External Error Modes (EEMs: what happened, e.g. closed wrong valve), or whether it also predicts Psychological Error Mechanisms (PEMs: how the operator failed internally, e.g. pattern recognition failure) and/or Performance Shaping Factors (PSF: situational factors that contribute to the likelihood of the error's occurrence, e.g. poor interface design; etc).

(v) *Usefulness*: the degree to which the technique can generate *error reduction mechanisms*, irrespective of whether these are based on analysis of root causes or not. This is judged as low (little concern of the technique with error reduction), moderate (suggesting that the technique is capable of error reduction), or high (meaning that error reduction is a primary focus of the approach, and that effective error reduction mechanisms will be generated either via detailed understanding of the error, or via sound engineering/design experience in devising alternative operational configurations of systems to avoid error opportunities). Usefulness also implicitly includes the criterion of *Diagnosticity*, here meaning the insight into the causes of the error, which allows (diagnostic) determination of error reduction measures.

(vi) *Resources* 1: likely resource usage in actually applying the technique, in terms of assessor/expert time. Resources were rated either as low, moderate or high, depending on the judged extent of time each technique would take to apply.

(vii) *Resources* 2: training required to use the system, i.e. the degree to which it is an expert's tool. This is simply rated as yes or no, since although this criterion could be rated as low, moderate and high these judgements would be very difficult to make without having used the systems comparatively.

(viii) *Resources* 3: the requirement for an expert panel or task-domain experts. This is rated simply on a yes/no basis.

(ix) *Documentability*: the degree to which the technique lends itself to auditable documentation. The techniques are rated as low (meaning that the way the technique is utilised is difficult to document), moderate (meaning that the technique provides sufficient documentation to be repeatable), or high (indicating that all assumptions etc. are recorded, and that in addition the documentation will be usable for future system operations and will greatly facilitate future periodic assessments).

(x) *Acceptability* 1: *PSA usage to date*. This is very difficult to judge, since so little has been published on usage of the techniques. A rating of low indicates that it appears that the technique has been developed but has only been used as a prototype. A rating of moderate indicates that it appears to have been used in a small number of assessments. A rating of high indicates that it has received extensive usage.

(xi) *Acceptability 2: Availability of technique.* This criterion indicates that the technique is either available (a rating of 'yes'), or else it is unavailable because it has been discontinued (in the case of PHECA), commercially related to one organisation (in the case of HRMS), or still at the prototype stage and not yet generally available (in the case of CES).

Additional criteria

Three additional criteria were developed for the purposes of this evaluation, based on judgement and a knowledge of HEI-for-PSA issues, in terms of what would be important indicators of a technique's overall utility for the NP & R PSA arena:

(xii) *HEI output quantifiability*: whether a special HRA quantification technique–HEI technique partnership exists between the HEI tool and e.g. Success Likelihood Index Method (SLIM: Embrey *et al.*, 1984), Absolute Probability Judgement (APJ) or Paired Comparisons (PC: see Kirwan, 1994), or THERP, or indeed whether the error forms developed are potentially beyond quantification at this stage.

(xiii) *Life cycle stage applicability*: the earliest life cycle stage at which the technique can probably be applied (concept; detailed design; commissioning; and existing/operational life cycle phases).

(xiv) *Primary objective of the technique*: the original purpose or function of the technique.

One notable missing criterion is that of validation evidence. This criterion is omitted because there have been so few validations (see however Whalley and Kirwan, 1989), so judgement on this criterion is not possible. It must also be stated at the outset that many of the judgements of the techniques made in the following tables are subjective and open to bias. Nevertheless it is useful to attempt such an evaluatory exercise to see which techniques appear ready for usage today and which appear promising for the future. In particular, where a technique is primarily *descriptive* in nature, rather than being *predictive*, this is noted in the table under this criterion by use of the letter [D].

Evaluation of HEI techniques

Each technique was therefore evaluated by the author[4] according to the criteria above in *Table 3*. Where there was insufficient information attainable, a question mark may occasionally appear in a cell in one of the table columns. Section 6 then discusses the results of the evaluations.

Summary evaluation of techniques

Although the techniques all share the common function of predicting errors, they have a wide range of functions. Nevertheless, some simple descriptive statistics can be used to draw insights from the above tables, and these are discussed below.

(i) Comprehensiveness: the number and proportion of techniques producing errors associated with each stage of behaviour or performance error type is as follows:

S (25: 66%); R (30: 79%); K (24: 63%); Rv (10: 26%); EOCs (7: 18%)

This is a fairly even split between S, R and K, with less emphasis on rule violations and EOCs. However, there is a trend towards the latter error types, and certainly new techniques would find it difficult to ignore these latter two error types if setting out to provide a HEI tool for HRA and PSA utilisation. Caution should be raised however, over the apparently high proportion of techniques dealing with knowledge-based behaviour, as many of these techniques are unavailable or highly specialised. There is therefore still room for the development of a more practicable and adaptable cognitive error prediction system.

(ii) The basis for the cognitive modelling simulations appears to be the following:

information processing (INTEROPS; CREWPRO; CAMEO-TAT; SYBORG);
symbolic processing theory (CES; COSIMO);
frame-based (COSIMO; CREWSIM (scripts));
blackboard architecture (CES; COSIMO)
Petri net (SEAMAID)

It has been noticeable in the recent five years that simulations have moved back towards information processing as a general framework. This may make them sit easier with Human Factors generally than a pure symbolic processing theoretic framework. The architectures adopted (frames; scripts; blackboards; Petri nets) are still maturing, and it is not clear which is best, though at a recent conference dealing with cognitive simulations (Hollnagel and Yoshikawa, 1996; 1998) it was noticeable that Petri nets were frequently in use as a practicable way of representing operator knowledge.

(iii) The degree to which techniques deal with all three error components is as follows:

EEMs (all); PEMs (13: 34%); PSF (23: 60%)

PSF are still seen as more useful in helping identify and reduce error likelihood, though techniques are increasingly considering PEMs also. What is clearly missing, however, are unified techniques wherein linkages between EEMs, PSF, and PEMs are specified (PHECA is a notable historical exception which attempted this). This appears to be an important, though difficult area requiring fundamental research rather than technique development. If such linkages could be developed, they would clarify and resolve many taxonomic problems that taxonomy (and technique) developers experience, allow better and more complete error reduction specification, and enable quantitative data to be collected in more robust and meaningful ways.

(iv) Expert's versus novice's tools: 71% of the techniques are tools requiring significant expertise in their use. This means that many of the tools will have limited application. Given that many tools do not seem to mature into field application tools, or if they do, their survival rate is not high, it would appear to be in developers'

[4]It should be noted that there is inevitably subjectivity (via the author) involved in the ratings of the techniques in the table, and due to paucity of information on many of the techniques, there may inevitably be some misleading representations and perhaps even some inaccuracies. Ideally the table would be developed by a group of HEI experts over a period of time, though there were not the resources to achieve this during this particular study.

Table 3 Evaluations of 38 HEI techniques

		THERP	INTENT	SRS-HRA	Murphy	CADA	TALENT	TAFEI	PHECA	HRMS
1.	Comprehensiveness	S,R	Rv,K,EOC	S,R,K, EOC	S,R,K	S,R,K	S,R	S,R,Rv,EOC	S,R,K,Rv	S,R,K,Rv
2.	Structuredness	Moderate	Low	Low	Moderate	Moderate	Low (judgement required)	Moderate	High	Moderate
3.1	Model-based?	Low	Low	N	High (SRK)	High (SRK/ Murphy)	Low	High	High (SRK)	High (SRK/ GEMS)
3.2	EEM/PEM/PSF	EEM	EEM/PSF	EEM PSF	EEM/PSF	EEM/PEM	EEM/PSF	EEM PSF	EEM/PEM/ PSF	EEM/PEM/
4.	Usefulness (ERA)	Low	Low	Low	High	High	Moderate	High (potentially)	High	High
5.1	Resources usage	Moderate	Low	Low	High	High	High	High	High	High
5.2	Experts tool	No	No	N	Yes	Yes	Yes	Moderate	Yes	Yes
5.3	Experts required	No	No	N	Yes	Yes	No	Yes	No	No
6.	Documentability	Moderate	Low: taxonomy	Low	High	High	Moderate	High	High	High
7.1	Usage in PSA?	High	Low (within SRS)	High	Low Moderate	Low/	Moderate	Low	Moderate	Moderate
7.2	Availability	Yes	Moderate	Y	Yes	Yes	Yes	Yes	No	Only to BNFL
8.	Quantifiable?	Yes (THERP)	APJ/rating/ SLIM	Y (various)	Yes (various)	Yes (various)	SLIM/ THERP	N/A	Yes (Various)	Yes (HRMS)
9.	Life cycle stage?	Concept	Operations	Detailed	Detailed	Detailed/ commission	Operating	Detailed	Detailed	Detailed
10.	Primary objective	HRA/PSA	PSA— quantitative inclusion of errors of intention	HEI for PSA	Design/ incident analysis [D]	Incident analysis/ HRA [D]	PSA, based on a detailed investigation of HF via task analysis	Error modes with equipment/ product design	HRA	HRA

		SRK	SHERPA	SCHEMA	TEACHER/ SIERRA	PHEA	GEMS	IMAS	CREAM	DYLAM
1.	Comprehensiveness	S,R,K	S,R	S,R	S,R	S, R	S,R,K,Rv	R,K	S,R,K,	K, R
2.	Structuredness	Moderate	High	Moderate	Moderate	Moderate	Low	Low	Moderate?	High
3.1	Model-based?	High (SRK)	High (SRK)	Moderate	Low	Moderate	High (GEMS)	Moderate	Moderate?	Y (fuzzy-set goal-means)
3.2	EEM/PEM/PSF	EEM/PEM/ PSF	EEM/PEM/ PSF	EEM/?/?	EEM/PSF	EEM (PSF?)	EEM/PEM/ PSF	EEM	EEM/PEM/ PSF	EEM/PEM
4.	Usefulness (ERA)	Moderate	High	Moderate?	High/Mod	Moderate	Moderate	Low	Moderate	Low
5.1	Resources Usage	High	Moderate	Moderate?	Moderate	Moderate	Moderate	High	High?	High
5.2	Experts tool	Yes	No	No?	Yes (at present)	No	Yes	Yes	Yes	Y
5.3	Experts required	No	No	No	Yes (at present)	No	No	Yes	No	Y

6.	Documentability	Low/ moderate	High	High?	High	High	Low	Moderate	Moderate	High
7.1	Usage in PSA?	Low	Moderate	Low	Low	Low	Low	Low	Low	Low
7.2	Availability	Yes chart form)	Yes (flow-	Yes?	Yes	Yes	Not as a formal technique	Low	Still being developed	N?
8.	Quantifiable?	Yes (various)	Yes (various)	SLIM	SLIM/APJ /N/A	SLIM/other	Yes (SLIM/APJ)	Yes (SLIM)	Not yet defined	Not developed
9. ?	Life cycle stage?	Detailed	Detailed	Detailed	Detailed	Detailed/ Commission	Detailed/ Commission	Detailed	Detailed	Operating
10.	Primary objective	Error categori- sation	HRA/ design assessment	PSA?; Design of operator support systems	Identifying potential low- consequence accidents	Errors leading to injury	Error categoris- ation	Diagnostic error	Risk assessment	Investigating cognitive behaviour; PSA?

		CES	INTEROPS	COSIMO	CREWSIM	CREWPRO	ADSA	CAMEO/ TAT	SYBORG	SEAMAID
1.	Comprehensiveness	R,K	K, R, S, Rv	K	R, K	R, K, S	K	S, R, K	S,R,K	S,R
2.	Structuredness	Moderate	High	High	High	High/Mod?	Moderate?	High	High	High
3.1	Model-based?	High	Y (Information processing)	High	Y (frames)	Y (information processing)	N	Y (information processing)	Information processing + emotions	Petri-net representation of knowledge
3.2	EEM/PEM/PSF	EEM/PEM/ limited PSF	EEM/PSF	EEM/PEM	EEM/PEM	EEM/PSF/ (PEM)	EEM/PSF	EEM/PSF	EEM/PEM/ PSF (emotions)	EEMs (via analyst)
4.	Usefulness (ERA)	Moderate	High	Low	Low	Moderate	Low	High	Moderate	High
5.1	Resources Usage	High	High	High	High	High	High?	High	High	High?
5.2	Experts tool	Yes	Y	Yes	Y	Y	Y	Y	Y	Yes
5.3	Experts required	No	Y?	Yes	Y (in setting up database)	Y?	Y?	N	Yes (under development)	N
6.	Documentability	High	High	High	High	High?	High	High	High	High
7.1	Usage in PSA?	No	N	Low	Low	Low	Low	Low	Low	Low
7.2	Availability	Low	N? (proprietary system)	Moderate	N?	N	N?	N	Still under development	Still under development
8.	Quantifiable?	Yes (APJ/ SLIM)	Not developed	Y?	Y?	Y?	Y	Y?	Yes?	Yes?
9.	Life cycle stage?	Detailed/ Commission	Detailed	Operating	Commission- ing	Existing	Commission- ing	Concept/ Detailed design	Commission- ing	Detailed design
10.	Primary objective	Diagnostic failure & procedures evaluation	Allocation of function	Cognitive modelling	Crew interactions	HRA	KBB errors	Task design & procedures	Evaluation of team performance	MMI design evaluation

Table 3 *Continued*

		HAZOP	CMA	TOPPE	PREDICT	EOCA	ATHEANA	HEMECA	SNEAK	PRMA	COMET	COGENT
1.	Comprehensiveness	S,R,Rv	R,K	R, comm's & team errors	Rv, K (S,R) EOCs	Rv,K,	EOCS, K, S, R and Rv?	S, R	S,R, EOCs	K, R, S, EOCs	EOCs	S,R,K
2.	Structuredness	Low	Low	Low	Low	Low	Moderate	Moderate	High	Moderate	Low	Low
3.1	Model-based?	Low	Low	N	Low	Low	Moderate	N	Moderate	N	Low	Moderate
3.2	EEM/PEM/PSF	EEM	EEM	EEM	EEM	EEM/PSF PEM?	EEM/PSF/	EEM/PSF	EEM/PSF	EEM/PSF	EEM/PSF	EEM
4.	Usefulness	High	Low	High	Moderate	Moderate	Moderate	High	Moderate	Moderate	Moderate/ low	Low
5.1	Resources usage	High	Moderate	High	High	High	High	High	High	High	Low	Low
5.2	Experts tool	No	No	Y	Yes	No	Y	Y	Yes	Y	Yes	Yes
5.3	Experts required	Yes	Yes	Y	Yes	Yes	Y	N	Yes	Y	Yes	Yes
6.	Documentability	High	Moderate	Moderate	High	Moderate	High	High	High	High	Moderate	Moderate
7.1	Usage in PSA?	Moderate	Moderate/ High	N	Low	Low	N/Low?	Low	Low	Moderate	Low	Low
7.2	Availability	Yes	Yes	Y	Yes	Yes	Prototype	Y	Yes	Y	Yes	Yes–under development
8.	Quantifiable?	Yes (various)	Yes (APJ/ PC/SLIM)	Y? (perhaps APJ)	APJ	Yes (APJ/SLIM)	Yes (APJ/SLIM)	Y (THERP; other)	Yes: various	Y (APJ; SLIM)	SLIM + THERP	SLIM/APJ?
9.	Life cycle stage?	Concept	Detailed	Commission-ing stage	Concept	Detailed/ Commiss.	Operations (existing)	Detailed	Operations	Operating	Operating	Detailed/ Operating
10.	Primary objective	PSA	Diagnostic failure	Procedural validation and team coordination	Violations of safety or design assumptions	Error of commission analysis	EOCs; other human events identification for PSA	Design	EOCs associated with HF deficiencies	KBB, EOCs in emergency response Scenarios	EOCs into PSA	Enriching cognitive content of PSAs

interests to develop tools that are more easily used, with little or no training (other than Human Factors or Reliability Engineering training, for example).

(v) 55% of the tools require experts as an input to the method application. This is roughly half, and demonstrates the difficulties in HEI of dealing with the richness of operational contexts, and of identifying errors in novel systems: hence the utilisation of human expert and context-relevant judgement. This usage of experts is likely to continue for some time in HEI, though it does inevitably raise the costs of using such tools, making them less attractive.

(vi) Only 24%, roughly a quarter of the techniques, appear to have actually had some usage in real PSAs. The other three quarters of techniques are either still under development or have not matured into field applicable tools of utility to industry. The 'hit rate' of technique developments is therefore 0.25. It is difficult to compare this to other fields, but it appears that there is certainly room for improvement, and perhaps room for more development and determination of what industry wants and what academia can offer, before proffering fledgling techniques into the market place.

(vii) The earliest life cycle application stage for the techniques appears to be as follows:

Concept (4: 10%)
Detailed (21: 55%)
Commissioning (4: 11%)
Existing (9: 24%)

This roughly bi-modal distribution suggests a predominance for techniques which support design evaluation, which is of benefit for Human Factors, and for techniques which deal with the majority of plants and installations that already exist. The fact that there are nevertheless some techniques applicable at the concept stage of the system development life cycle, however, is also good both for engineering design and reliability, and Human Factors, since the sooner errors are identified, the sooner and more cheaply and effectively they can be eradicated.

(viii) The primary objectives of the techniques are as follows:

HRA/PSA/Performance & error modelling (30: 79%)
Design aspects (4: 11%)
Accident analysis (2: 5%)
Low consequence HRA (2: 5%)

Clearly the dominant reason for developing these techniques is HRA/PSA. Although only a small number of techniques have low risk/consequence HRA as their priority, it should be noted that many of the 'larger-HRA' HEI techniques will also deal with low-risk HRA issues (e.g. HAZOP; SHERPA; etc. and other techniques have been applied in diverse environments from manufacturing, to mail delivery, for example). Whilst design aspects are relatively low as a primary purpose, that is because there are many other techniques suited specifically for this purpose. The same is also true of accident and incident analysis. However, an observation is that there seems to be an unhelpful division between accident analysis techniques and taxonomies, and HEI tools. Since both are looking at the same thing, one looking into the past, and one into the future, there should be a synergy between these two areas of investigation, and perhaps even a more formal coherence between taxonomies from the two areas. Error taxonomies should be Janis-like[1]: they should be capable of looking equally into the past and the future.

Technique development requirements for HRA/PSA

For the practitioner, the major question is which techniques can be used now, or in the near future, and which techniques in contrast are no longer being pursued or are simply unavailable. For the more general reader, and for the developers of future techniques, the techniques that appear to be worth pursuing in the short term, or in the longer term, are of most interest, for it is the future techniques that may be able to expand the functionality of what HEI can achieve, making HEI more comprehensive and incisive, and also potentially broadening its application base from the sole concern of HRA, into other Human Factors domains. Therefore, based on the evaluations in the foregoing tables, in particular whether the techniques are available now, and their special features, the techniques can be usefully grouped into five categories, as follows:

- tools which are usable now,
- tools which could be adapted or will be available in the short term to be useful for HRA/PSA,
- tools which are unavailable, or whose availability is highly restricted,
- tools whose focus is not on high risk complex system HRA/PSA, or which have been superseded and are no longer in use in HRA/PSA,
- tools requiring longer term development.

The evaluation placing the thirty-eight techniques into these categories is depicted in *Figure 1*.

The arrows highlight the progression of development of the techniques to reach practicable tools, but also hint that many prototypes fail to reach true maturity and develop as usable and used tools in real PSA/HRA applications. HEI appears to be a field rich in ideas and yet relatively low in fruition.

The main techniques which are available in the public domain, and could be of general practical use now in PSA, are therefore as follows:

- THERP
- HUMAN-HAZOP
- SHERPA (and derivatives)
- CMA/FSMA
- PRMA
- EOCA (an EOC-related form of HAZOP)
- SRS-HRA (a contextual taxonomy: a more generic form could be developed for each industry type-the JHEDI technique [Kirwan, 1990] offers a UK-based alternative)
- SRK flowchart and error forms
- HRMS (in its proprietary company, along with the JHEDI approach)

[1]Janis was a mythological character who could look both into the past and the future, and is the root of the name of the month of January.

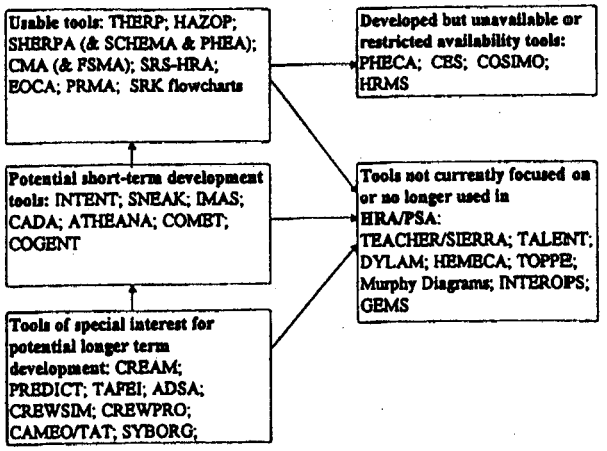


Figure 1 Current status of techniques

These techniques can be amalgamated and transformed into a HRA/PSA assessor's toolkit, in the short term. The question is then one of comprehensiveness: are all relevant error types being addressed? If the answer is yes, then there is no point in developing new approaches, but if the answer is 'no', then identification of poorly addressable error types can target areas of human involvement most urgently needing development. In the early part of this paper, a number of error types were addressed, namely the following:

- slips and lapses (also called action execution errors),
- cognitive errors: diagnostic and decision-making errors,
- maintenance errors and latent failures,
- errors of commission,
- rule violations (and management errors/safety culture failures).

Currently management failures (safety culture failures) are left outside of the PSA process, i.e. they are dealt with in a qualitative way (see Kennedy and Kirwan, 1995, 1996, 1998). However, a small number of the techniques have been specifically concerned with communication failures and team cohesion problems (e.g. CREWSIM and CREWPRO; TOPPE). Such potential errors are important in certain scenarios where operation and control is *distributed* and there is reliance on co-ordination of interacting groups and individuals, and on oral (but not necessarily face-to-face) communication. In cognitive psychology terms, this is the area of *Distributed Decision-Making*, and it is of concern to risk assessment, as there is large potential for error impact in such scenarios. Since communication errors may be due to slips or lapses, or more cognitive misinterpretations, and may also lie dormant for some time, and since team co-ordination and cohesion problems probably require separate task and error analysis, it may be useful to separate these two error forms from the others when deciding whether they are adequately catered for by current techniques. This has been done in *Table 4*, which shows a matrix of techniques versus types of errors (note that maintenance and latent failures are subsumed under skill and rule based errors).

Table 4 also highlights the fact that although skill and rule-based slips and lapses are adequately catered for with "mature" techniques, other areas of human involvement (particularly rule violations and teams/communication errors) are not. Furthermore, cognitive errors are addressed by a number of techniques, but few of these cognitive simulations reach sufficient maturity to become useful and used in a PSA environment. There is clearly much further development to fully address human involvement in complex systems. However, this must be set against the fact that most errors will probably be of the skill and rule type. Nevertheless, this counter-point can itself be countered, since such routine errors are probably not the ones that will lead to a reactor core melt or the loss of an offshore platform. Further research and development is therefore clearly needed.

Concluding comments

Adequacy of HEI tools

A general review of the evaluations in the above tables suggests that there is no single technique available at present which would be optimal on all the qualitative criteria: there is therefore no clear 'best' technique currently available. Many of the techniques are still not highly structured, suggesting that the majority of the techniques are still seen as aiding the assessor rather than being fully prescriptive and model driven, and implying that HEI is still an art rather than a well-defined science (even the simulation models appear to require significant assessor input and judgement in their application as HEI tools). The second paper in this pair of papers takes the viewpoint that a framework approach might resolve some of the problems and weak areas currently facing practitioners, who cannot wait for new techniques to be developed and mature. Such an approach places a number of approaches or taxonomies into a framework which aims to address all error types. This then becomes a structured 'toolkit' for the practitioner, and can be used in a resource-flexible way. The second paper will therefore exemplify this concept with a practicable prototype approach developed for the UK NP & R industry.

Longer term development requirements

The major trend at present appears to be EOC and cognitive error identification (in the USA these two often being merged, but in the UK EOCs represent rule violations and slips/lapses leading to event states not associated with the goal of the task under investigation). The other minor trend is for more model-based approaches, i.e. models rather than judgmental or incident-based taxonomies. It is also noticeable, however, that a number of methods attract large investment resources and are then discontinued, shortly before they might become useful in the PSA arena, this apparently being the fate of CES and perhaps COSIMO, and potentially the future fate of CREWSIM and INTEROPS, as far as PSA goes. This may also be a function of the apparent proliferation of prototypical models, which vary in their model-make-up (i.e. information processing versus symbolic processing, etc.). Such diversity of models is of course academically interesting and important, but it might be better to concentrate on one and reach the stage where it can begin to be practicable for PSA, than to 'get it right first time' from a purely theoretical standpoint. The former could be achieved by developing a model with the sole aim of being usable in PSA. Such a development life-cycle

Table 4 HEI tools of potential practical use for UK NP and R PSA

Error type	Skill and Rule-based errors	Cognitive errors	Violations and errors of commission	Comm'n and team errors
Technique(s) available, appropriate, and practicable	THERP; HAZOP; SHERPA; (HRMS); SNEAK; PHEA; SCHEMA; SRS-HRA	CMA; PRMA	HAZOP EOCA	
Aspects of other techniques which could aid HEI	SRK flowchart; Murphy diagrams framework; HRMS ERA module;	GEMS taxonomy; SRK	PREDICT keywords	
Tools which could be developed or tailored in the near future (e.g. within a year)	TAFEI	IMAS; CREAM	ATHEANA INTENT SNEAK PREDICT TAFEI	TOPPE; communication error taxonomy
Tools which require longer term development (e.g. 3–5 y)	SEAMAID; CAMEO/TAT	CREWSIM; ADSA; GEMS	SYBORG	CREWSIM; CREWPRO; SYBORG

would probably be of the order of three years if an existing prototype was selected (e.g. CREWSIM), and 5+ years if starting 'from scratch'. The benefits of such a model, however, are worth the investment, since the increase in cognitive modelling resolution and validity is potentially large when compared to desk-based approaches such as CMA/FSMA. It would also yield more effective error reduction guidance for cognitively demanding scenarios.

Another significant and recent advance is the modelling of team interactions, including communication errors and confidence in one's fellow crew members and in incoming messages and information. CREWSIM seems to be the leader in the field with respect to crew interactions, though this aspect of the model has not been under development for some time (Siu and Dang, 1994). Nevertheless, communication in emergency and routine situations is often a significant contributor/cause of real events, and this development is encouraging in that perhaps communication error analysis (including distributed decision making) is within sight, at least in the long term.

Perhaps the most novel addition to HEI tools has been SYBORG, which aims to consider emotional states of operators. Whether this approach succeeds or not, this development certainly broadens the theoretical potential of HEI to address the full spectrum of errors and their causes and, as with other technique developments already cited, reflects the trend towards analysing the context in which the operators think and act in work systems.

The development of soundly based EOC approaches is less clearly visible. Whereas the theory for cognitive error is relatively developed (symbolic processing and information-processing models, coupled with SRK and GEMS error forms), that for rule violations is not. The potential for any slip to lead to an EOC will mean that slip based EOCA will therefore continue to be resource intensive. Therefore what is needed is development of more theories of rule violations and their associated error forms and mechanisms, and more tractable methods on the lines of SNEAK and/or PREDICT, or even TAFEI (if the state space transitions can in some way be computed and screened in terms of relative likelihood).

More fundamentally, development of understanding of the inter-relationships between EEMs, PEMs, and PSF, in different contextual environments, would greatly advance the utility of HEI. Furthermore, such theoretical and empirical development work would place HEI (and HRA) on a more scientific footing. Since human error is likely to continue to play a dominant causative role in risk in society, a scientific approach towards predicting and managing human error should be able to make a substantial contribution towards helping to reduce accident frequency and severity. There is still a long way to go, but it is a worthwhile path.

Acknowledgements

This work was sponsored by the Industry Management Committee for the UK Generic Nuclear Safety Research budget under contract HF/GNSR/22 from April '94 until March '96. The author would like to thank the following for their comments at various meetings: David Whitfield (NII), Ray Hughes (NE), Mike Gray (HSE), Kay Ryder (BNFL), and in particular for the comments and support of Keith Rea (BNFL) who was also the Project Officer. The author would also like to thank the two referees, who made a series of valuable contributions to the scope and direction of the paper. The contents of this paper are, however, the opinions of the author, and do not necessarily reflect those of the above-mentioned personnel nor their respective companies.

References

Amendola, A., Reina, G. and Ciceri, F. (1985) 'Dynamic simulation of man-machine interaction in incident control' *Paper in IFAC Man-Machine Systems Conference*, Varese, Italy

Baber, C. and Stanton, N. (1991) 'Task analysis for error identification: towards a methodology for identifying human error' In **Lovesey, E. J.** (ed.) Contemporary Ergonomics, Taylor and Francis, London: pp. 67–71.

Beith, B. H., Vail, R. E. and Drake, M. (1991) 'The team operations performance and procedure evaluation (TOPPE) technique as a method for evaluating emergency operating procedures' In *Human Factors Society Annual Conference Proceedings*, pp. 405–411

Bersini, U., Cacciabue, P. C. and Mancini, G. (1988) 'Cognitive modelling: a basic complement of human reliability analysis' *Reliability Engng System Safety* **22**, 107–128.

Blackman, H. S. (1991) 'Modelling the influence of errors of commission on success probability' *Proceedings of the 35th Annual Human Factors Society Meeting*, San Francisco, 2–6 September pp 1085–1089

Cacciabue, P. C., Decortis, F., Drozdowicz, B., Masson, M. and Nordvik, J.-P. (1992) 'COSIMO: a cognitive simulation model of human decision making and behaviour in accident management of complex plants' *IEEE Trans on Systems Man Cybernet* **22**, (5), 1058–1074

Comer, P. J., Fitt, F. S. and Ostebo, R. (1986) 'A driller's HAZOP Method' Society of Petroleum Engineers (SPE) 15867

Cooper, S. E., Ramey-Smith, A. M., Wreathall, J., Parry, G. W., Bley, D. C., Luckas, W. J., Taylor, J. H. and Barriere, M. T. (1996) 'A technique for human error analysis (ATHEANA) – Technical Basis and Method Description' Nureg/CR-6350. US Nuclear Regulatory Commission, Washington DC 20555

Cox, S. J. and Tait, N. R. S. (1991) 'Reliability, Safety and Risk Management' Oxford: Butterworth–Heinemann

Dang, V., Huang, Y., Siu, N. and Carroll, J. (1993) 'Analysing cognitive errors using a dynamic crew-simulation model. pp. 520–525

Dang, V. and Siu, N. (1994) 'Simulating operator cognition for risk analysis: current models and CREWSIM' *Paper in PSAM-II Proceedings*, San Diego, California, 20–25 March pp. 066-7–066-13

Embrey, D. E., Humphreys, P. C., Rosa, E. A., Kirwan, B. and Rea, K. (1984) 'SLIM-MAUD: An approach to assessing human error probabilities using structured expert judgement'. NUREG/CR-3518, US Nuclear Regulatory Commission, Washington, DC 20555

Embrey, D. E. (1986a) 'SHERPA: a systematic human error reduction and prediction approach'. Paper Presented at the International Topical Meeting on Advances in Human Factors in Nuclear Power Systems, Knoxville, Tennessee

Embrey, D. E. (1986b) 'Approaches to aiding and training operators' diagnoses in abnormal situations' Chem Ind 454–459

Embrey, D. E. (1993) 'Techniques for auditing and reducing risk from human error' In *Ergonomics in the Process Industries*, IChemE North-Western Branch Papers No. 5, Rugby, IChemE

Embrey, D. E., Kontogiannis, T. and Green, M. (1994) 'Preventing human error in process safety' Centre for Chemical Process Safety (CCPS), American Inst of Chem Eng, New York: CCPS

Fujita, Y., Sakuda, H. and Yanagisawa, I. (1994) 'Human reliability analysis using simulated human model'. In *PSAM-II Proceedings*, San Diego, California, 20–25, March, pp. 060–13–060-18

Gall, W. (1988) *Error Analysis–SRD Human Reliability Course Notes*, UKAEA, Culcheth, Cheshire

Green, A. E. (1983) *Safety systems reliability* Chichester, Wiley

Gertman, D. I. (1991) 'INTENT: a method for calculating HEP estimates for decision–based errors' *Proceedings of the 35th Annual Human Factors Meeting*, San Francisco, 2–6 September pp. 1090–1094

Gertman, D. I. (1993) 'Representing cognitive activities and errors in HRA trees' Reliability Engng System Safety **39**, 25–34

Gertman, D. I., Haney, L. N. and Ostrom, L. T. (1994) 'Evidence of the need to model errors of commission in risk assessments for varied environments' In *PSAM-II Proceedings*, San Diego, California, 20–25 March, pp. 005-15–005-20

Gibson, J. J. (1979) *The Ecological Approach to Visual Perception.* Houghton-Mifflin New York.

Hahn, H. A. and deVries, J. A. (1991) 'Identification of human errors of commission using Sneak Analysis' *Proceedings of the Human Factors Society 35th Annual Meeting*, San Francisco, 2–6 September pp. 1080–1084

Hasegawa, N. and Yoshimura, S. (1996) '*Emotions and human errors in* a control room of a nuclear power plant–for the development of an operator team behaviour model with emotional function' In **Yoshikawa, H. and Hollnagal, E.** (Eds.), Cognitive Science Engineering in Process Control. Kyoto, 12–15 November pp. 151–158

Hollnagel, E. (1993) *Human Reliability Analysis*: Context and Control. Academic Press London

Hollnagel, E. and Embrey, D. E. (1994) *Human Reliability Assessment in Safety Assessments*: Addressing Cognitive and Action Errors. HRA Ltd, Parbold, Lancashire

Hollnagel, E. (1996) Personal communication

Hsueh, K-S., Soth, L. and Mosleh, A. (1994) 'A simulation study of errors of commission in nuclear power accidents' In *PSAM-II Proceedings*, San Diego, California, 20–25 March, pp. 066-1–066-6

Kennedy, R. and Kirwan, B. (1995) The failure mechanisms of safety culture. *IAEA Conference on Safety Culture*, Vienna, April 1995

Kennedy, R. and Kirwan, B. (1998) 'A Safety Culture HAZOP approach'. *Safety Science* (In press)

Kirwan, B., Embrey, D. E. and Rea, K. (1988) The human reliability assessors guide, Report RTS 88/95Q, NCSR, UKAEA, Culcheth, Cheshire

Kirwan, B. and James, N. J. (1989) 'The development of a human reliability assessment system for the management of human error in complex systems' in Reliability 89, Brighton Metropole, June

Kirwan, B. (1990) 'A resources flexible approach to human reliability assessment for PSA' *Safety and Reliability Symposium*, Altrincham, September 1990 Elsevier London

Kirwan, B. (1992a) 'Human error identification in human reliability assessment. Part 1: overview of approaches' Applied Ergonomics, **23** (5), 299–318

Kirwan, B. (1992b) 'Human error identification in HRA. Part 2: Detailed comparison of techniques'. *Applied Ergonomics*, **23** (6), 371–381

Kirwan, B. and Ainsworth L. A. (eds.) (1992) A *Guide to Task Analysis*, Taylor & Francis London

Kirwan, B. (1993) 'A Human error analysis toolkit for complex systems' Paper presented at the 4th Cognitive Science Approaches to Process Control Conference, 25–27 August, Copenhagen, Denmark, pp 151–199

Kirwan, B. (1994) *A Guide to Practical HRA* Taylor and Francis, London

Kirwan, B., Scannali, S. and Robinson, L. (1995) *Practical HRA in PSA–a case study. European Safety and Reliability Conference, ESREL '95*, Bournemouth, 26–28 June Institute of Quality Assurance. pp 675–693

Kirwan, B., Scannali, S. and Robinson, L. (1996) 'Applied HRA: a case study' *Applied Ergonomics*, **27** (5) 289–302.

Kirwan, B. (1996) 'The requirements of cognitive simulations for human reliability and probabilistic safety assessments' Paper presented at the First Cognitive Science Engineering in Process Control Conference, Kyoto, Japan. 12–15 November

Kirwan, B. and Hollnagel, E. (1998) 'The requirements of cognitive simulations for human reliability and probabilistic safety assessment' in **Hollnagel, E. and Yoshikawa, H.** (eds.) *Cognitive Systems Engineering in Process Control*, (in press)

Kletz, T. (1974) *HAZOP and HAZAN – Notes on the Identification and Assessment of Hazards.* Institute of Chemical Engineers, Rugby

Livingston, A. D., Wright, M. S. and Embrey, D. E. (1992) 'The application of task analysis and human error analysis to the development of operating instructions in a batch chemical process plant' in **Kragt, H.** (ed.) *Enhancing Industrial Performance: Experiences of integrating the Human Factor*, Taylor & Francis, London pp. 280–295

Nakagawa, T., Nakatani, Y., Yoshikawa, H., Takahashi, M., Furuta, T., and Hasegawa, A. (1996) 'Development of simulation based evaluation support system for man–machine interface design' in **Cacciabue, P. C. and Papazoglou, I. A.** (eds.) *SEAMAID system. PSAM III*, Springer, London pp. 1185–1190.

Parry, G. W. (1994) 'A procedure for the analysis of errors of commission in a PSA' in *PSAM-II Proceedings*, March 20–25 San Diego, California

Pew, R. W., Miller, D. C. and Feehrer C. S. (1981) 'Evaluation of proposed control room improvements through analysis of critical operator decisions' NP 1982 Palo Alto, CA Electric Power Research Institute

Potash, L. *et al* (1981) 'Experience in integrating the operator contribution in the PSA of actual operating plants' in Proceedings of the ANS/ENS Topical Meeting on PSA, American Nuclear Society, New York

Rasmussen, J. *et al* (1981) 'Classification system for reporting events involving human malfunction' Riso–M–2240, DK–4000, Riso National Laboratories, Denmark

Reason, J. T. (1987a) 'A framework for classifying errors in **Rasmussen, J., Duncan, K. D. and Leplat, J.** (eds) *New Technology and Human Error* Chichester, Wiley

Reason, J. T. (1987b) 'Generic error modelling system' in **Rasmussen, J., Duncan, K. and Leplat, J.** (eds) A Cognitive Framework for Locating Common Human Error Forms in New Technology and Human Error, Chichester, Wiley

Reason, J. T. and Embrey D. E. (1985) 'Human factors principles relevant to the modelling of human errors' Report for the European Atomic Energy Community

Reason, J. T. (1990) Human Error Cambridge University Press, Cambridge, UK

Ryan, T. (1988) 'A task analysis-linked approach for integrating the human factor in reliability assessments of nuclear power plants' *Reliability Engng and Sys Safety*, 22, 219–234.

Schryver, J. C. and Knee, H. E. (1987) 'Integrated operator plant process modelling and decision support for allocation of function' Paper Presented at the 31st Human Factors Society Conference, pp. 815–819

Schryver, J. C. (1991) 'Operator model–based design and evaluation of advanced systems: computational models' *Proceedings of the IEEE Fourth Conference on Human Factors and Power Plants* pp. 121–127

Shen, S-H., Smidts, C. and Mosleh, A. (1994) 'Elements of a model for operator problem solving and decision making in abnormal conditions'. In *PSAM-II Proceedings*, San Diego, California, 20–25 March pp. 060-7–060-12

Siu, N. and Dang, V. (1994) Personal communication

Stanton, N. (1994) 'Task analysis for error identification: a review of the technique' Paper Presented at the Risk Assessment and Risk Reduction Conference, Aston University, 22nd March

Swain, A. D. and Guttmann, H. E. (1983) – 'A handbook of human reliability analysis with emphasis on nuclear power plant applications' *USNRC-Nureg/CR-1278* Washington DC.

Takano, K., Sasou, K. and Yoshimura, S. (1995) 'Simulation system for behaviour of an operating group' 14th European Annual Conference on Decision–Making and Manuall Control, June 1995, Delft, Netherlands

Vail, R. E., Eide, H. C., Benhardt, H. C., Held, J. E. and Olsen, L. M. (1994) 'Human error model development for the Savannah River Site non-reactor facilities.' in *PSAM-II*, San Diego, California, 20–25 March, pp. 073-13–073-18

Whalley, S. P. (1988) – Minimising the cause of human error' in **G. P. Libberton** *10th Advances in Reliability Technology Symposium* Elsevier London

Whalley, S. P. and Kirwan, B. (1989) 'An evaluation of five human error identification techniques' Paper Presented at the 5th International Loss Prevention Symposium, Oslo, June 1989

Whittingham, R. B. and Reed, J. (1989) 'Identification and reduction of critical human error using a FMEA approach' Paper Presented at Reliability 89, Brighton Metropole, June, 1989 pp. 5A/1/1–5A/1/8

Williams, J. C. (1986) 'HEART – a proposed method for assessing and reducing human error' *9th Advances in Reliability Technology Symposium*, University of Bradford

Williams, J. C. and Munley, G. A. (1992) "Human error identification – a new approach' Paper Presented at PSA/PRA, Safety and Risk Assessment IBC, London, 3/4 December

Woods, D. D., Roth, E. M. and Pople, M. (1987) 'An artificial intelligence based cognitive model for human performance assessment' USNRC NUREG/CR-4862. Washington DC

Woods, D. D., Pople, H. E. and Roth, E. M. (1990) The cognitive environment simulation as a tool for modelling human performance and reliability' *USNRC, Nureg/CR-5213* Washington DC